GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:21:32 ; Search time 48 Seconds
(without alignments)
52.978 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	100.0	9	4	AAB48919	Aab48919 Generic S
2	45	100.0	9	4	AAB48917	Aab48917 SH2 domai
3	45	100.0	9	4	AAB48922	Aab48922 SH2 domai
4	45	100.0	9	5	ABG68582	Abg68582 Peptide G
5	45	100.0	10	4	AAB48923	Aab48923 SH2 domai
6	45	100.0	10	4	AAB48920	Aab48920 SH2 domai
7	45	100.0	10	4	AAB48926	Aab48926 SH2 domai
8	45	100.0	10	4	AAB48921	Aab48921 SH2 domai
9	45	100.0	10	4	AAB48928	Aab48928 SH2 domai
10	45	100.0	11	2	AAW46897	Aaw46897 G1C-S pep
11	45	100.0	11	2	AAW46896	Aaw46896 Non-phosp
12	45	100.0	11	5	ABG68419	Abg68419 G1 peptid
13	45	100.0	11	5	ABG68583	Abg68583 Peptide G
14	45	100.0	26	4	AAB48932	Aab48932 SH2 domai
15	45	100.0	26	4	AAB48933	Aab48933 SH2 domai
16	37	82.2	11	2	AAW46899	Aaw46899 Non-phosp
17	37	82.2	919	2	AAW63117	Aaw63117 Human ade
18	36	80.0	11	2	AAW46898	Aaw46898 Non-phosp
19	36	80.0	20	2	AAR49328	Aar49328 Influenza
20	36	80.0	20	2	AAW54715	Aaw54715 Peptide f
21	36	80.0	244	2	AAW80804	Aaw80804 Amino aci
22	36	80.0	244	2	AAW95053	Aaw95053 Myrotheci
23	36	80.0	448	6	ABU19327	Abu19327 Protein e
24	36	80.0	562	2	AAR63588	Aar63588 Full leng
25	36	80.0	562	5	AAE23111	Aae23111 Influenza
26	36	80.0	921	6	AAO23317	Aao23317 Rhesus mo
27	36	80.0	931	6	AAO23313	Aao23313 Cynomolgu
28	35	77.8	9	2	AAY10382	Aay10382 T cell ep
29	35	77.8	9	5	ABG80064	Abg80064 MHC class
30	35	77.8	9	7	ADC35620	Adc35620 Influenza
31	35	77.8	84	6	ADA08462	Ada08462 Human AFA
32	35	77.8	86	6	ADA08458	Ada08458 Chicken A
33	35	77.8	86	6	ADA08461	Ada08461 Avian AFA
34	35	77.8	362	2	AAY13465	Aay13465 Peptide S
35	35	77.8	634	4	AAB93557	Aab93557 Human pro
36	35	77.8	815	6	ADA08456	Ada08456 Chicken A
37	34	75.6	293	5	ABG93283	Abg93283 C. albica
38	34	75.6	3542	4	AAB62142	Aab62142 P. falcip
39	33	73.3	10	4	AAB48925	Aab48925 SH2 domai
40	33	73.3	10	4	AAB48927	Aab48927 SH2 domai
41	33	73.3	38	2	AAR58364	Aar58364 TSAR bind
42	33	73.3	310	6	ABM68832	Abm68832 Photorhab
43	33	73.3	434	4	AAU33491	Aau33491 Enterococ
44	33	73.3	448	4	AAU35058	Aau35058 Enterococ
45	33	73.3	448	6	ABU14570	Abu14570 Protein e

ALIGNMENTS

RESULT 1
AAB48919
ID AAB48919 standard; peptide; 9 AA.
XX
AC AAB48919;
XX
DT 16-MAR-2001 (first entry)
XX
DE Generic SH2 domain cyclic peptide inhibitor, SEQ ID NO:3.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .9
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety, thereby
FT cyclising the peptide"
FT Misc-difference 1
FT /note= "Any naturally or non-naturally occurring amino
FT acid except Glu"
FT Modified-site 9
FT /note= "C-terminal amide"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Disclosure; Page 5; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding

```
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence is a generic
CC   representation of a cyclic peptide of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 2
AAB48917
ID   AAB48917 standard; peptide; 9 AA.
XX
AC   AAB48917;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain cyclic peptide inhibitor, SEQ ID NO:1.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .9
FT                   /note= "The nitrogen atoms of the N-terminus and the C-
FT                   terminal amide are joined via a bridging moiety, thereby
FT                   cyclising the peptide"
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
FT   Modified-site   9
FT                   /note= "C-terminal amide"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Claim 1; Page 21; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a cyclic
CC   peptide of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 45;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 3
AAB48922
ID   AAB48922 standard; peptide; 9 AA.
XX
AC   AAB48922;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:7.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid; the nitrogen atom of
FT                   the N-terminus is joined to a ClCH2C(O) moiety"
FT   Modified-site   9
FT                   /note= "The carbon atom of the C-terminus is joined to a
FT                   C(CH2SH)C(O)NH2 moiety"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
```

PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 1; Page 13; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 45;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9
```

RESULT 4
ABG68582
ID ABG68582 standard; peptide; 9 AA.
XX
AC ABG68582;
XX
DT 07-OCT-2002 (first entry)
XX
DE Peptide G1TE #1.
XX
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; G1;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
PN WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI Krag DN, Pero SC, Oligino L;
XX
DR WPI; 2002-547451/58.
XX
PT Treatment or prophylaxis of a subject having a disorder characterized by
PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Fig 9B; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal
CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC placenta, colon, ovary, testes and lung. The present sequence is a G1
CC peptide (not defined) or derivative which is used to illustrate the
CC possible structures of cyclic Grb7 antagonists
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 45;  DB 5;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.4e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||||||
Db          1 ELYENVGMY 9
```

RESULT 5
AAB48923
ID AAB48923 standard; peptide; 10 AA.
XX
AC AAB48923;
XX
DT 16-MAR-2001 (first entry)
XX
DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:8.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .10
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety, thereby
FT cyclising the peptide"
FT Modified-site 1
FT /label= Aad
FT Modified-site 10
FT /note= "C-terminal amide"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX

DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 2; Page 13; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 45; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
||||||||
Db 1 XLYENVGMY 9

RESULT 6
AAB48920
ID AAB48920 standard; peptide; 10 AA.
XX
AC AAB48920;
XX
DT 16-MAR-2001 (first entry)
XX
DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:4.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .10
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety C(O)-CH2-
FT S-CH2-CHC(O)NH2, thereby cyclising the peptide"
FT Modified-site 1
FT /note= "Gamma-carboxyglutamic acid"
FT Modified-site 10
FT /note= "C-terminal amide"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 1; Page 12; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 45; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
||||||||
Db 1 XLYENVGMY 9

RESULT 7
AAB48926
ID AAB48926 standard; peptide; 10 AA.
XX
AC AAB48926;
XX
DT 16-MAR-2001 (first entry)
XX
DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:11.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; linear precursor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 10
FT /label= Nle
FT /note= "C-terminal amide, joined to a solid matrix"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.

XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 4; Page 14; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 45; DB 4; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.014;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|||||||
Db 1 ELYENVGMY 9

RESULT 8
AAB48921
ID AAB48921 standard; peptide; 10 AA.
XX
AC AAB48921;
XX
DT 16-MAR-2001 (first entry)
XX
DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:5.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; linear precursor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Gamma-carboxyglutamic acid"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 1; Page 12; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 45; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
|||||||||
Db 1 XLYENVGMY 9

RESULT 9
AAB48928
ID AAB48928 standard; peptide; 10 AA.
XX
AC AAB48928;
XX
DT 16-MAR-2001 (first entry)
XX
DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:14.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; linear precursor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 10
FT /label= Aad
FT /note= "C-terminal amide, joined to a solid matrix"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 5; Page 15; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
SQ Sequence 10 AA;

```
  Query Match             100.0%;  Score 45;  DB 4;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.014;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||||||| 
Db          1 ELYENVGMY 9
```

RESULT 10
AAW46897
ID AAW46897 standard; peptide; 11 AA.
XX
AC AAW46897;
XX
DT 19-JUN-1998 (first entry)
XX
DE G1C-S peptide.
XX
KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
KW signal transduction protein; non-phosphorylated; inhibition; treatment;
KW hyper-proliferative disease; human cancer.
XX
OS Unidentified.
XX
PN WO9802176-A1.
XX
PD 22-JAN-1998.
XX
PF 16-JUL-1997; 97WO-US012501.
XX
PR 16-JUL-1996; 96US-0021858P.
XX
PA (GEOU) UNIV GEORGETOWN.
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI King CR, Sastry L, Krag D, Oligino L;
XX
DR WPI; 1998-110340/10.
XX
PT Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
PT transducing protein - at least as well as natural phosphorylated target,
PT particularly from treatment of cancer.
XX
PS Disclosure; Page 18; 39pp; English.
XX
CC The present sequence represents a peptide designated G1C-S. This peptide
CC is essentially the same as a non-phosphorylated peptide, G1, that is
CC capable of binding to the src homology 2 (SH2) domain of Grb2, except
CC that the terminal Cys residues of G1 are replaced with Ser residues. Grb2
CC is a signal transduction protein. The binding affinity of the present
CC peptide with Grb2 was tested, and it was demonstrated that the disulphide
CC bond of G1 may be important. The G1 peptide binds to the SH2 domain of
CC Grb2 with affinity similar to, or greater than, that of a SHC
CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
CC that has not been modified by phosphate or similar charged group. The G1
CC peptide is used to inhibit a signal transduction process that involves
CC binding of a phosphorylated protein or peptide to the SH2 domain of a
CC signal transduction protein, particularly Grb2. It is used specifically
CC for treatment of hyper-proliferative diseases, especially human cancer
XX
SQ Sequence 11 AA;

```
  Query Match             100.0%;  Score 45;  DB 2;  Length 11;
  Best Local Similarity   88.9%;  Pred. No. 0.015;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||||||| 
Db          2 ELYENVGMY 10
```

RESULT 11
AAW46896
ID AAW46896 standard; peptide; 11 AA.
XX
AC AAW46896;
XX
DT 19-JUN-1998 (first entry)
XX
DE Non-phosphorylated peptide which binds to the SH2 domain of Grb2.
XX
KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
KW signal transduction protein; non-phosphorylated; inhibition; treatment;
KW hyper-proliferative disease; human cancer; cyclic.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1. .11
XX
PN WO9802176-A1.
XX
PD 22-JAN-1998.
XX
PF 16-JUL-1997; 97WO-US012501.
XX
PR 16-JUL-1996; 96US-0021858P.
XX
PA (GEOU) UNIV GEORGETOWN.
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI King CR, Sastry L, Krag D, Oligino L;
XX
DR WPI; 1998-110340/10.
XX
PT Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
PT transducing protein - at least as well as natural phosphorylated target,
PT particularly from treatment of cancer.
XX

PS Claim 9; Page 17; 39pp; English.
XX
CC The present sequence represents non-phosphorylated peptide, G1, that is
CC capable of binding to the src homology 2 (SH2) domain of Grb2. Grb2 is a
CC signal transduction protein. The G1 peptide binds to the SH2 domain of
CC Grb2 with affinity similar to, or greater than, that of a SHC
CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
CC that has not been modified by phosphate or similar charged group. The G1
CC peptide is used to inhibit a signal transduction process that involves
CC binding of a phosphorylated protein or peptide to the SH2 domain of a
CC signal transduction protein, particularly Grb2. It is used specifically
CC for treatment of hyper-proliferative diseases, especially human cancer
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 45; DB 2; Length 11;
Best Local Similarity 88.9%; Pred. No. 0.015;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|||||||
Db 2 ELYENVGMY 10

RESULT 12
ABG68419
ID ABG68419 standard; peptide; 11 AA.
XX
AC ABG68419;
XX
DT 07-OCT-2002 (first entry)
XX
DE G1 peptide.
XX
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; G1;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
PN WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI Krag DN, Pero SC, Oligino L;
XX
DR WPI; 2002-547451/58.
XX
PT Treatment or prophylaxis of a subject having a disorder characterized by
PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Page 102; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal
CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC placenta, colon, ovary, testes and lung. The present sequence is a G1
CC peptide (not defined) or derivative which is used to illustrate the
CC possible structures of cyclic Grb7 antagonists
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 45; DB 5; Length 11;
Best Local Similarity 88.9%; Pred. No. 0.015;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|||||||
Db 2 ELYENVGMY 10

RESULT 13
ABG68583
ID ABG68583 standard; peptide; 11 AA.
XX
AC ABG68583;
XX
DT 07-OCT-2002 (first entry)
XX
DE Peptide G1TE #2.
XX
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; G1;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
PN WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI Krag DN, Pero SC, Oligino L;
XX
DR WPI; 2002-547451/58.
XX
PT Treatment or prophylaxis of a subject having a disorder characterized by
PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Fig 9C; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal

```
CC   interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC   cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC   gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC   placenta, colon, ovary, testes and lung. The present sequence is a G1
CC   peptide (not defined) or derivative which is used to illustrate the
CC   possible structures of cyclic Grb7 antagonists
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 45;  DB 5;  Length 11;
  Best Local Similarity   88.9%;  Pred. No. 0.015;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||||||||
Db          2 ELYENVGMY 10


RESULT 14
AAB48932
ID   AAB48932 standard; peptide; 26 AA.
XX
AC   AAB48932;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:18.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 12; Page 19; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a linear
CC   precursor of a peptide of the invention
XX
SQ   Sequence 26 AA;

  Query Match             100.0%;  Score 45;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 15
AAB48933
ID   AAB48933 standard; peptide; 26 AA.
XX
AC   AAB48933;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain cyclic peptide inhibitor, SEQ ID NO:19.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .10
FT                   /note= "The nitrogen atom of the N-terminus and the Cys
FT                   10 sidechain are joined via a bridging moiety, thereby
FT                   cyclising part of the peptide"
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 12; Page 20; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
```

CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 45; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9
```

Search completed: July 15, 2004, 07:28:49
Job time : 51 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:26:37 ; Search time 14.5 Seconds
(without alignments)
32.044 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	86.7	566	2	US-08-272-255-8	Sequence 8, Appli
2	39	86.7	566	5	PCT-US95-08565-8	Sequence 8, Appli
3	37	82.2	919	2	US-08-788-674-4	Sequence 4, Appli
4	36	80.0	19	4	US-09-376-343-3	Sequence 3, Appli
5	36	80.0	20	2	US-08-480-190-38	Sequence 38, Appl
6	36	80.0	20	2	US-08-488-379-38	Sequence 38, Appl
7	36	80.0	20	4	US-08-475-399A-38	Sequence 38, Appl
8	36	80.0	20	5	PCT-US93-07545-38	Sequence 38, Appl
9	36	80.0	244	3	US-09-003-287-6	Sequence 6, Appli
10	36	80.0	244	3	US-09-003-287-8	Sequence 8, Appli
11	36	80.0	244	3	US-09-518-988-2	Sequence 2, Appli
12	35	77.8	9	1	US-08-146-145-6	Sequence 6, Appli
13	35	77.8	362	2	US-09-080-897-6	Sequence 6, Appli
14	35	77.8	362	3	US-09-323-735-6	Sequence 6, Appli
15	33	73.3	38	1	US-08-176-500-22	Sequence 22, Appl
16	33	73.3	38	1	US-08-471-052A-22	Sequence 22, Appl
17	33	73.3	38	1	US-08-189-331-22	Sequence 22, Appl
18	33	73.3	38	2	US-08-471-939-22	Sequence 22, Appl
19	33	73.3	38	2	US-08-471-800-22	Sequence 22, Appl
20	33	73.3	38	2	US-08-471-068-22	Sequence 22, Appl
21	33	73.3	245	4	US-09-134-000C-3547	Sequence 3547, Ap
22	33	73.3	310	4	US-09-252-991A-27339	Sequence 27339, A
23	33	73.3	693	4	US-09-376-343-2	Sequence 2, Appli
24	32	71.1	15	1	US-08-176-500-31	Sequence 31, Appl
25	32	71.1	15	1	US-08-471-052A-31	Sequence 31, Appl
26	32	71.1	15	1	US-08-189-331-31	Sequence 31, Appl
27	32	71.1	15	2	US-08-471-939-31	Sequence 31, Appl
28	32	71.1	15	2	US-08-471-800-31	Sequence 31, Appl
29	32	71.1	15	2	US-08-488-161-20	Sequence 20, Appl
30	32	71.1	15	2	US-08-471-068-31	Sequence 31, Appl
31	32	71.1	15	3	US-09-273-685-20	Sequence 20, Appl
32	32	71.1	15	5	PCT-US95-11934-20	Sequence 20, Appl
33	32	71.1	445	4	US-09-489-039A-13869	Sequence 13869, A
34	32	71.1	461	2	US-08-527-227A-7	Sequence 7, Appli
35	32	71.1	485	4	US-09-543-681A-4935	Sequence 4935, Ap
36	32	71.1	487	1	US-08-249-112-4	Sequence 4, Appli
37	32	71.1	487	5	PCT-US95-06556-4	Sequence 4, Appli
38	32	71.1	593	1	US-08-202-389-12	Sequence 12, Appl
39	32	71.1	593	1	US-08-018-129-5	Sequence 5, Appli
40	32	71.1	593	2	US-08-448-250-5	Sequence 5, Appli
41	32	71.1	593	4	US-09-282-257-5	Sequence 5, Appli
42	32	71.1	605	2	US-08-752-307B-8	Sequence 8, Appli
43	32	71.1	605	4	US-09-707-802-8	Sequence 8, Appli
44	32	71.1	605	4	US-09-991-326-8	Sequence 8, Appli
45	32	71.1	671	3	US-09-132-118-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-272-255-8
; Sequence 8, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/272,255
; FILING DATE: 08-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-272-255-8

Query Match 86.7%; Score 39; DB 2; Length 566;
Best Local Similarity 66.7%; Pred. No. 8.2;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMY 9
              :|||:|||:|
Db         87 RLYDNVGLY 95
```

```
RESULT 2
PCT-US95-08565-8
; Sequence 8, Application PC/TUS9508565
;  GENERAL INFORMATION:
;     APPLICANT:  Cashmore, Anthony R.
;     APPLICANT:  Ahmad, Margaret
;     APPLICANT:  Lin, Chentao
;     TITLE OF INVENTION:  Blue Light Photoreceptors and Methods of
;     TITLE OF INVENTION:  Using the Same
;     NUMBER OF SEQUENCES:  22
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz & Norris
;       STREET:  One Liberty Place, 46th floor
;       CITY:  Philadelphia
;       STATE:  PA
;       COUNTRY:  USA
;       ZIP:  19103
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/08565
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/272,255
;       FILING DATE:  08-JUL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Leary Ph.D., Kathryn
;       REGISTRATION NUMBER:  36,317
;       REFERENCE/DOCKET NUMBER:  UPN-1795
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-3100
;       TELEFAX:  (215) 568-3439
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  566 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US95-08565-8

  Query Match             86.7%;  Score 39;  DB 5;  Length 566;
  Best Local Similarity   66.7%;  Pred. No. 8.2;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:|||:|
Db         87 RLYDNVGLY 95


RESULT 3
US-08-788-674-4
; Sequence 4, Application US/08788674
; Patent No. 5922315
;  GENERAL INFORMATION:
;     APPLICANT:  Roy, Soumitra
;     TITLE OF INVENTION:  Adenoviruses Having Altered
;     TITLE OF INVENTION:  Hexon Proteins
;     NUMBER OF SEQUENCES:  7
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Carella, Byrne, Bain,
;       ADDRESSEE:  Gilfillan, Cecchi, Stewart &
;       ADDRESSEE:  Olstein
;       STREET:  6 Becker Farm Road
;       CITY:  Roseland
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07068
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch diskette
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  Word Perfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/788,674
;       FILING DATE:  24-JAN-1997
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Olstein, Elliot M.
;       REGISTRATION NUMBER:  24,025
;       REFERENCE/DOCKET NUMBER:  271010-363
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  973-994-1700
;       TELEFAX:  973-994-1744
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  919 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FEATURE:
;       NAME/KEY:  predicted hexon protein sequence
;       NAME/KEY:  for human Adenovirus 12
US-08-788-674-4

  Query Match             82.2%;  Score 37;  DB 2;  Length 919;
  Best Local Similarity   66.7%;  Pred. No. 36;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        439 FLYSNVGLY 447


RESULT 4
US-09-376-343-3
; Sequence 3, Application US/09376343
; Patent No. 6506592
; GENERAL INFORMATION:
;  APPLICANT: Blum, Paul H.
;  TITLE OF INVENTION: Hyperthermophilic Alpha-Glucosidase Gene and Its Use
;  FILE REFERENCE: N1231-200
;  CURRENT APPLICATION NUMBER: US/09/376,343
;  CURRENT FILING DATE: 1999-08-18
;  EARLIER APPLICATION NUMBER: 60/096,860
;  EARLIER FILING DATE: 1998-08-18
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Sulfolobus solfataricus
US-09-376-343-3

  Query Match             80.0%;  Score 36;  DB 4;  Length 19;
  Best Local Similarity   55.6%;  Pred. No. 0.74;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              ::|||:|:|
Db          5 KIYENLGVY 13


RESULT 5
US-08-480-190-38
; Sequence 38, Application US/08480190
```

```
; Patent No. 5827516
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,190
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-480-190-38

  Query Match             80.0%;  Score 36;  DB 2;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 0.78;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db          2 TLYQNVGTY 10


RESULT 6
US-08-488-379-38
; Sequence 38, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,379
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-488-379-38

  Query Match             80.0%;  Score 36;  DB 2;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 0.78;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db          2 TLYQNVGTY 10


RESULT 7
US-08-475-399A-38
; Sequence 38, Application US/08475399A
; Patent No. 6509033
;  GENERAL INFORMATION:
;    APPLICANT:  Urban, Robert G.
;    APPLICANT:  Chicz, Roman M.
;    APPLICANT:  Vignali, Dario A.A.
;    APPLICANT:  Hedley, Mary L.
;    APPLICANT:  Stern, Lawrence J.
;    APPLICANT:  Strominger, Jack L.
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  276
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,399A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  15-JUN-1993
```

```
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  11-AUG-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fraser, Janis K.
;       REGISTRATION NUMBER:  34,819
;       REFERENCE/DOCKET NUMBER:  00246/168003
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617/542-507
;       TELEFAX:  617/542-890
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  38:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-475-399A-38

  Query Match             80.0%;  Score 36;  DB 4;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 0.78;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db          2 TLYQNVGTY 10


RESULT 8
PCT-US93-07545-38
; Sequence 38, Application PC/TUS9307545
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  273
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07545
;      FILING DATE:  19930811
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US93-07545-38

  Query Match             80.0%;  Score 36;  DB 5;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 0.78;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db          2 TLYQNVGTY 10


RESULT 9
US-09-003-287-6
; Sequence 6, Application US/09003287
; Patent No. 6096947
; GENERAL INFORMATION:
;  APPLICANT: Jayne, Susan
;  APPLICANT: Barbour, Eric
;  APPLICANT: Meyer, Terry
;  TITLE OF INVENTION: METHODS FOR IMPROVING TRANSFORMATION EFFICIENCY
;  FILE REFERENCE: moPAT_moCAH
;  CURRENT APPLICATION NUMBER: US/09/003,287
;  CURRENT FILING DATE: 1998-01-06
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Myrothecium verrucaria
US-09-003-287-6

  Query Match             80.0%;  Score 36;  DB 3;  Length 244;
  Best Local Similarity   66.7%;  Pred. No. 13;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        169 TLYDNVGAY 177


RESULT 10
US-09-003-287-8
; Sequence 8, Application US/09003287
; Patent No. 6096947
; GENERAL INFORMATION:
;  APPLICANT: Jayne, Susan
;  APPLICANT: Barbour, Eric
;  APPLICANT: Meyer, Terry
;  TITLE OF INVENTION: METHODS FOR IMPROVING TRANSFORMATION EFFICIENCY
;  FILE REFERENCE: moPAT_moCAH
;  CURRENT APPLICATION NUMBER: US/09/003,287
;  CURRENT FILING DATE: 1998-01-06
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Myrothecium verrucaria
US-09-003-287-8

  Query Match             80.0%;  Score 36;  DB 3;  Length 244;
  Best Local Similarity   66.7%;  Pred. No. 13;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        169 TLYDNVGAY 177


RESULT 11
US-09-518-988-2
; Sequence 2, Application US/09518988
; Patent No. 6268547
;  GENERAL INFORMATION:
;    APPLICANT:  Weeks, James T.
```

; TITLE OF INVENTION: TRANSFORMATION OF WHEAT WITH THE
; TITLE OF INVENTION: CYANAMIDE HYDRATASE GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nancy J. Parsons
; STREET: 800 Buchanan St.
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/518,988
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/873,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parsons, Nancy J.
; REGISTRATION NUMBER: 40,364
; REFERENCE/DOCKET NUMBER: 0177.95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-5731
; TELEFAX: (510) 559-5736
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-518-988-2

Query Match 80.0%; Score 36; DB 3; Length 244;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:||:||| |
Db 169 TLYDNVGAY 177

RESULT 12
US-08-146-145-6
; Sequence 6, Application US/08146145
; Patent No. 5747269
; GENERAL INFORMATION:
; APPLICANT: Rammensee, Hans-Georg
; APPLICANT: Falk, Kirsten
; APPLICANT: R tzschke, Olaf
; APPLICANT: Stevanovic, Stefan
; APPLICANT: Jung, G nther
; TITLE OF INVENTION: DETERMINATION OF PEPTIDE MOTIFS ON MHC
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,145
; FILING DATE: 17-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitts, Monica C.
; REGISTRATION NUMBER: 36,105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-145-6

Query Match 77.8%; Score 35; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LYENVGMY 9
||:||| |
Db 1 LYQNVGTY 8

RESULT 13
US-09-080-897-6
; Sequence 6, Application US/09080897
; Patent No. 5985574
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Welcsh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-080-897-6

Query Match 77.8%; Score 35; DB 2; Length 362;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMY 9
              :||||:| |
Db        247 KLYENLGEY 255


RESULT 14
US-09-323-735-6
; Sequence 6, Application US/09323735
; Patent No. 6197932
;  GENERAL INFORMATION:
;    APPLICANT:  King, Mary-Claire
;    APPLICANT:  Lynch, Eric D.
;    APPLICANT:  Lee, Ming
;    APPLICANT:  Morrow, Jan E.
;    APPLICANT:  Welcsh, Piri L.
;    APPLICANT:  Leon, Pedro E.
;    TITLE OF INVENTION:  Modulators of Actin
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  75 DENISE DRIVE
;      CITY:  HILLSBOROUGH
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/323,735
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/080,897
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UW97-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 343-4341
;      TELEFAX:  (650) 343-4342
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  362 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-323-735-6

  Query Match             77.8%;  Score 35;  DB 3;  Length 362;
  Best Local Similarity   66.7%;  Pred. No. 32;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||||:| |
Db        247 KLYENLGEY 255


RESULT 15
US-08-176-500-22
; Sequence 22, Application US/08176500
; Patent No. 5498538
;  GENERAL INFORMATION:
;    APPLICANT:  Kay, B. K.
;    APPLICANT:  Fowlkes, D. M.
;    TITLE OF INVENTION:  Totally Synthetic Affinity Reagents
;    NUMBER OF SEQUENCES:  141
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/176,500
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/013,416
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-143
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-176-500-22

  Query Match             73.3%;  Score 33;  DB 1;  Length 38;
  Best Local Similarity   66.7%;  Pred. No. 6.4;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| | |||
Db         14 LLYANPGMY 22


Search completed: July 15, 2004, 07:31:18
Job time : 15.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:27:08 ; Search time 40 Seconds
(without alignments)
70.326 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOM8.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOM8.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	100.0	9	10	US-09-998-350-1	Sequence 1, Appli
2	45	100.0	9	10	US-09-998-350-3	Sequence 3, Appli
3	45	100.0	9	10	US-09-998-350-7	Sequence 7, Appli
4	45	100.0	10	10	US-09-998-350-4	Sequence 4, Appli
5	45	100.0	10	10	US-09-998-350-5	Sequence 5, Appli
6	45	100.0	10	10	US-09-998-350-6	Sequence 6, Appli
7	45	100.0	10	10	US-09-998-350-8	Sequence 8, Appli
8	45	100.0	10	10	US-09-998-350-11	Sequence 11, Appl
9	45	100.0	10	10	US-09-998-350-14	Sequence 14, Appl
10	45	100.0	11	14	US-10-013-815-32	Sequence 32, Appl
11	45	100.0	26	10	US-09-998-350-18	Sequence 18, Appl
12	45	100.0	26	10	US-09-998-350-19	Sequence 19, Appl
13	36	80.0	244	15	US-10-392-301-33	Sequence 33, Appl
14	36	80.0	448	12	US-10-282-122A-47251	Sequence 47251, A
15	35	77.8	9	12	US-10-367-580-48	Sequence 48, Appl
16	35	77.8	9	12	US-10-367-593-48	Sequence 48, Appl
17	35	77.8	9	12	US-10-367-594-48	Sequence 48, Appl
18	35	77.8	9	12	US-10-367-654-48	Sequence 48, Appl
19	35	77.8	9	12	US-10-367-658-48	Sequence 48, Appl
20	35	77.8	9	12	US-10-367-668-48	Sequence 48, Appl
21	35	77.8	9	16	US-10-367-674-48	Sequence 48, Appl
22	35	77.8	9	16	US-10-777-053-366	Sequence 366, App
23	35	77.8	9	16	US-10-777-053-943	Sequence 943, App
24	35	77.8	9	16	US-10-777-053-958	Sequence 958, App
25	35	77.8	79	14	US-10-246-354-7	Sequence 7, Appli
26	35	77.8	84	14	US-10-246-354-10	Sequence 10, Appl
27	35	77.8	86	14	US-10-246-354-6	Sequence 6, Appli
28	35	77.8	168	12	US-10-424-599-170035	Sequence 170035,
29	35	77.8	815	14	US-10-246-354-3	Sequence 3, Appli
30	35	77.8	1096	16	US-10-408-765A-747	Sequence 747, App
31	34	75.6	1234	15	US-10-369-493-13287	Sequence 13287, A
32	34	75.6	3542	12	US-10-087-013-2	Sequence 2, Appli
33	33	73.3	10	10	US-09-998-350-10	Sequence 10, Appl
34	33	73.3	10	10	US-09-998-350-12	Sequence 12, Appl
35	33	73.3	10	10	US-09-998-350-13	Sequence 13, Appl
36	33	73.3	78	12	US-10-424-599-219681	Sequence 219681,
37	33	73.3	134	16	US-10-437-963-168439	Sequence 168439,
38	33	73.3	162	12	US-10-424-599-205104	Sequence 205104,
39	33	73.3	306	15	US-10-369-493-1088	Sequence 1088, Ap
40	33	73.3	434	9	US-09-815-242-4987	Sequence 4987, Ap
41	33	73.3	448	9	US-09-815-242-10651	Sequence 10651, A
42	33	73.3	448	12	US-10-282-122A-42494	Sequence 42494, A
43	33	73.3	693	14	US-10-228-063-5	Sequence 5, Appli
44	33	73.3	712	14	US-10-228-063-27	Sequence 27, Appl
45	33	73.3	718	14	US-10-228-063-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-09-998-350-1
; Sequence 1, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is an amide, i.e. C(O)NH
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (1)..(9)
;   OTHER INFORMATION: Xaa (Gla) and Tyr at position 9 are bridged together, making this
;   OTHER INFORMATION: peptide cyclic
US-09-998-350-1

  Query Match            100.0%;  Score 45;  DB 10;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 2
US-09-998-350-3
; Sequence 3, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa is any amino acid other than Glu
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(9)
;   OTHER INFORMATION: Xaa and Tyr at position 9 are bridged together, making this pepti
;   OTHER INFORMATION: de cyclic
US-09-998-350-3

  Query Match            100.0%;  Score 45;  DB 10;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 3
US-09-998-350-7
; Sequence 7, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa has a ClCH2C(O)- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 has a -C(CH2SH)C(O)NH2 group attached
US-09-998-350-7

  Query Match            100.0%;  Score 45;  DB 10;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 4
US-09-998-350-4
; Sequence 4, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys at position 10 is an amide, i.e., C(O)NH
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Xaa (Gla) and Cys are bridged together, making this peptide cycli
;   OTHER INFORMATION: c
US-09-998-350-4

  Query Match           100.0%;  Score 45;  DB 10;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 5
US-09-998-350-5
; Sequence 5, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
US-09-998-350-5

  Query Match           100.0%;  Score 45;  DB 10;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 6
US-09-998-350-6
; Sequence 6, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla(OtBu)2, which is di- tert-butoxy-gamma-carboxy-L-glutam
;   OTHER INFORMATION: ic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: Tyr at position 3 is modified to Tyr(tBu), which is tert-butyl-ty
;   OTHER INFORMATION: rosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is is trytyl-asp
;   OTHER INFORMATION: aragine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is modified to Tyr(tBu), which is tert-butyl-ty
;   OTHER INFORMATION: rosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys at position 10 is modified to Cys(Trt), which is trytyl-cyste
;   OTHER INFORMATION: ine, and Cys(Trt) is connected to a resin
US-09-998-350-6

  Query Match           100.0%;  Score 45;  DB 10;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 7
US-09-998-350-8
; Sequence 8, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
```

```
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Adi, which is alpha-amino-adipic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa has a CH2CO- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Xaa (Adi) and Cys are bridged together, making this peptide cycli
;   OTHER INFORMATION: c
US-09-998-350-8

  Query Match           100.0%;  Score 45;  DB 10;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 0.034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 8
US-09-998-350-11
; Sequence 11, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: Tyr at position 3 is modified to Tyr(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: tyrosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trityl-aspara
;   OTHER INFORMATION: gine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: tyrosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa = Nle, which is norleucine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa is an amide and is attached to a resin
US-09-998-350-11

  Query Match           100.0%;  Score 45;  DB 10;  Length 10;
  Best Local Similarity  88.9%;  Pred. No. 0.034;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||||||:
Db          1 ELYENVGMY 9


RESULT 9
US-09-998-350-14
; Sequence 14, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trytyl-aspara
;   OTHER INFORMATION: gine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: tyrosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa = Adi(OA1), which is allyloxy-alpha-amino-adipic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa is an amide, i.e., C(O)NH
US-09-998-350-14

  Query Match             100.0%;  Score 45;  DB 10;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.034;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||||||||
Db          1 ELYENVGMY 9


RESULT 10
US-10-013-815-32
; Sequence 32, Application US/10013815
; Publication No. US20030105000A1
; GENERAL INFORMATION:
;  APPLICANT: Pero, Stephanie
;  APPLICANT:  Krag, David
;  APPLICANT:  Oligino, Lyn
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING GRB7
;  FILE REFERENCE: V0139/7048 (HCL/MAT)
;  CURRENT APPLICATION NUMBER: US/10/013,815
;  CURRENT FILING DATE:  2001-11-05
;  PRIOR APPLICATION NUMBER: US 60/245,755
;  PRIOR FILING DATE: 2000-11-03
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20030105000A1-phosphorylated peptide with YEN motif
US-10-013-815-32

  Query Match             100.0%;  Score 45;  DB 14;  Length 11;
  Best Local Similarity   88.9%;  Pred. No. 0.037;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||||||||
Db          2 ELYENVGMY 10


RESULT 11
US-09-998-350-18
; Sequence 18, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
US-09-998-350-18

  Query Match             100.0%;  Score 45;  DB 10;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.095;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 12
US-09-998-350-19
; Sequence 19, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa (G1a) has a CH2CO- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
US-09-998-350-19

  Query Match            100.0%;  Score 45;  DB 10;  Length 26;
  Best Local Similarity  100.0%;  Pred. No. 0.095;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              |||||||||
Db          1 XLYENVGMY 9


RESULT 13
US-10-392-301-33
; Sequence 33, Application US/10392301
; Publication No. US20040003434A1
; GENERAL INFORMATION:
;  APPLICANT: WEEKS, J. TROY
;  APPLICANT:  ROMMENS, CAIUS
;  TITLE OF INVENTION: REFINED PLANT TRANSFORMATION
;  FILE REFERENCE: 058951/0164
;  CURRENT APPLICATION NUMBER: US/10/392,301
;  CURRENT FILING DATE:  2003-03-20
;  PRIOR APPLICATION NUMBER: 60/365,527
;  PRIOR FILING DATE: 2002-03-20
;  PRIOR APPLICATION NUMBER: 60/377,597
;  PRIOR FILING DATE: 2002-05-06
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Myrothecium verrucaria
US-10-392-301-33

  Query Match            80.0%;  Score 36;  DB 15;  Length 244;
  Best Local Similarity  66.7%;  Pred. No. 59;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        169 TLYDNVGAY 177


RESULT 14
US-10-282-122A-47251
; Sequence 47251, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47251
;   LENGTH: 448
;   TYPE: PRT
;   ORGANISM: Borrelia burgdorferi
US-10-282-122A-47251

  Query Match            80.0%;  Score 36;  DB 12;  Length 448;
  Best Local Similarity  55.6%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|||::|:|
Db        336 LLYEDIGLY 344


RESULT 15
US-10-367-580-48
; Sequence 48, Application US/10367580
; Publication No. US20040071720A1
; GENERAL INFORMATION:
;  APPLICANT: Rothman, James E.
;  APPLICANT:  Hartl, F. Ulrich
;  APPLICANT:  Hoe, Mee H.
;  APPLICANT:  Houghton, Alan
;  APPLICANT:  Takechi, Yoshizumi
;  APPLICANT:  Mayhew, Mark
;  TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
;  FILE REFERENCE: 11746/461061
;  CURRENT APPLICATION NUMBER: US/10/367,580
;  CURRENT FILING DATE:  2003-02-14
;  PRIOR APPLICATION NUMBER: US 09/794,832
;  PRIOR FILING DATE: 2001-02-27
;  PRIOR APPLICATION NUMBER: US 09/011,645
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: PCT/US96/13363
;  PRIOR FILING DATE: 1996-08-16
;  PRIOR APPLICATION NUMBER: US 60/002,490
;  PRIOR FILING DATE: 1995-08-18
;  PRIOR APPLICATION NUMBER: US 60/002,479
;  PRIOR FILING DATE: 1995-08-18
;  NUMBER OF SEQ ID NOS: 349
;  SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 48
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide
US-10-367-580-48

  Query Match            77.8%;  Score 35;  DB 12;  Length 9;
```

Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 LYENVGMY 9
              ||:||| |
Db          1 LYQNVGTY 8
```

Search completed: July 15, 2004, 07:32:49
Job time : 40 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:23:22 ; Search time 11.5 Seconds
(without alignments)
75.280 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	86.7	565	2	S67298	deoxyribodipyrimid
2	37	82.2	468	2	S37217	hexon protein - hu
3	37	82.2	526	1	VGVNSG	spike glycoprotein
4	37	82.2	919	2	S33942	hexon protein - hu
5	36	80.0	20	2	PL0161	hemagglutinin - In
6	36	80.0	244	2	A39365	cyanamide hydratas
7	36	80.0	447	2	S39296	hexon protein - hu
8	36	80.0	448	1	F70190	probable diphospha
9	36	80.0	562	1	HMIV2	hemagglutinin prec
10	36	80.0	936	2	S57637	hexon protein - hu
11	35	77.8	29	2	B81136	hypothetical prote
12	35	77.8	34	2	H81883	hypothetical prote
13	35	77.8	150	2	A55883	actin-filament-ass
14	34	75.6	511	2	A99574	ABC transporter at
15	34	75.6	1249	2	A56511	myosin I myoA - Em
16	33	73.3	99	2	S44632	f22b7.3 protein -
17	33	73.3	306	2	D64497	aspartate carbamoy
18	33	73.3	309	2	F83044	nitrate-inducible
19	33	73.3	312	2	A64042	formate dehydrogen
20	33	73.3	313	2	A57499	N5-(carboxyethyl)o
21	33	73.3	332	2	T33774	hypothetical prote
22	33	73.3	352	2	D72264	hypothetical prote
23	33	73.3	354	2	E97128	magnesium and coba
24	33	73.3	389	2	B81380	hypothetical prote
25	33	73.3	439	2	G88103	protein W10G11.17
26	33	73.3	512	1	VGIVTH	envelope glycoprot
27	33	73.3	661	2	S49901	coat protein gp1 -
28	33	73.3	688	2	T33708	hypothetical prote
29	33	73.3	693	2	H90486	alpha-glucosidase
30	33	73.3	700	2	T20550	hypothetical prote
31	33	73.3	739	2	AI1876	hypothetical prote
32	33	73.3	852	2	T33824	hypothetical prote
33	32	71.1	149	2	S67188	hypothetical prote
34	32	71.1	221	2	E64400	conserved hypothet
35	32	71.1	224	2	H89847	hypothetical prote
36	32	71.1	231	2	H85138	hypothetical prote
37	32	71.1	263	2	AI0471	probable exported
38	32	71.1	307	2	T34973	5,10-methylenetetr
39	32	71.1	336	2	T30459	hypothetical prote
40	32	71.1	434	2	S50865	avermectin-sensiti
41	32	71.1	437	2	A64891	coenzyme F390 synt
42	32	71.1	450	1	DCCHO	ornithine decarbox
43	32	71.1	455	1	DCHYOC	ornithine decarbox
44	32	71.1	460	2	A43563	ornithine decarbox
45	32	71.1	461	1	DCHUO	ornithine decarbox

ALIGNMENTS

RESULT 1
S67298
deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O6771; protein YOR386w
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C;Accession: S67298; A23964; A24046
R;Delius, H.; Hebling, U.; Hofmann, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67261
A;Accession: S67298
A;Molecule type: DNA
A;Residues: 1-565 <DEL>
A;Cross-references: EMBL:Z75294; NID:g1420830; PIDN:CAA99718.1; PID:g1420831; MIPS:YOR38
A;Experimental source: strain S288C
R;Yasui, A.; Langeveld, S.A.
Gene 36, 349-355, 1985
A;Title: Homology between the photoreactivation genes of Saccharomyces cerevisiae and Esc
A;Reference number: A23964; MUID:86083177; PMID:3000886
A;Accession: A23964
A;Molecule type: DNA
A;Residues: 1-76,'A',78-164,'S',166-168,'T',170-199,'S',201-350,'R',352-364,'E',366-472,'
A;Cross-references: EMBL:M11578; NID:g172169; PIDN:AAA34875.1; PID:g172170
R;Sancar, G.B.
Nucleic Acids Res. 13, 8231-8246, 1985
A;Title: Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of the PHR1 phot
A;Reference number: A24046; MUID:86067229; PMID:3906569
A;Accession: A24046
A;Molecule type: DNA
A;Residues: 1-565 <SAN>
A;Cross-references: EMBL:X03183; NID:g4175; PIDN:CAA26944.1; PID:g4176
C;Genetics:
A;Gene: SGD:PHR1
A;Cross-references: SGD:S0005913; MIPS:YOR386w
A;Map position: 15R
C;Superfamily: deoxyribodipyrimidine photo-lyase
C;Keywords: carbon-carbon lyase

Query Match 86.7%; Score 39; DB 2; Length 565;
Best Local Similarity 66.7%; Pred. No. 4.6;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMY 9
              :||:|||:|
Db         86 RLYDNVGLY 94
```

RESULT 2
S37217
hexon protein - human adenovirus 31 (fragment)
C;Species: Mastadenovirus h31 (human adenovirus 31)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

```
C;Accession: S37217
R;Pring-Akerblom, P.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37213
A;Accession: S37217
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-468 <PRI>
A;Cross-references: EMBL:X74661; NID:g402765; PIDN:CAA52725.1; PID:g402766
C;Superfamily: adenovirus hexon protein

  Query Match             82.2%;  Score 37;  DB 2;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 9.7;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        341 FLYSNVGLY 349


RESULT 3
VGVNSG
spike glycoprotein G precursor - sigma virus
C;Species: sigma virus
A;Note: host Drosophila melanogaster
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C;Accession: A27150
R;Teninges, D.; Bras-Herreng, F.
J. Gen. Virol. 68, 2625-2638, 1987
A;Title: Rhabdovirus sigma, the hereditary CO-2 sensitivity agent of Drosophila: nucleot
A;Reference number: A27150; MUID:88034947; PMID:2822842
A;Accession: A27150
A;Molecule type: genomic RNA
A;Residues: 1-526 <TEN>
A;Cross-references: GB:X06171; NID:g61818; PIDN:CAA29536.1; PID:g61819
C;Genetics:
A;Gene: G
A;Cross-references: FlyBase:FBgn0015809
C;Superfamily: rhabdovirus spike glycoprotein G
C;Keywords: glycoprotein; spike protein; transmembrane protein
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-526/Product: spike glycoprotein G #status predicted <SGG>
F;499-515/Domain: transmembrane #status predicted <TMN>
F;32,445,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             82.2%;  Score 37;  DB 1;  Length 526;
  Best Local Similarity   66.7%;  Pred. No. 11;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||::||||
Db        350 VLYQSVGMY 358


RESULT 4
S33942
hexon protein - human adenovirus 12
N;Alternate names: late protein 2
C;Species: Mastadenovirus h12 (human adenovirus 12)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S33942
R;Sprengel, J.
submitted to the EMBL Data Library, June 1993
A;Reference number: S33928
A;Accession: S33942
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-919 <SPR>
A;Cross-references: EMBL:X73487; NID:g313361; PIDN:CAA51891.1; PID:g313376
C;Superfamily: adenovirus hexon protein

  Query Match             82.2%;  Score 37;  DB 2;  Length 919;
  Best Local Similarity   66.7%;  Pred. No. 21;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        439 FLYSNVGLY 447


RESULT 5
PL0161
hemagglutinin - Influenza H2N2 (fragment)
C;Species: influenza H2N2
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-May-1997
C;Accession: PL0161
R;Sweetser, M.T.; Braciale, V.L.; Braciale, T.J.
J. Exp. Med. 170, 1357-1368, 1989
A;Title: Class I major histocompatibility complex-restricted T lymphocyte recognition of
A;Reference number: PL0161; MUID:90010790; PMID:2477491
A;Accession: PL0161
A;Molecule type: mRNA
A;Residues: 1-20 <SWE>
A;Experimental source: strain A/JAP/305/57
C;Comment: This protein plays a major role in initiation of infection and in the pathoger
C;Superfamily: influenza virus hemagglutinin
C;Keywords: hemagglutinin
F;1-20/Region: immunodominant site recognized by T-lymphocytes

  Query Match             80.0%;  Score 36;  DB 2;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 0.45;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db          2 TLYQNVGTY 10


RESULT 6
A39365
cyanamide hydratase (EC 4.2.1.69) - fungus (Myrothecium verrucaria)
C;Species: Myrothecium verrucaria
C;Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 15-Sep-2000
C;Accession: A39365
R;Maier-Greiner, U.H.; Obermaier-Skrobranek, B.M.M.; Estermaier, L.M.; Kammerloher, W.; F
R.
Proc. Natl. Acad. Sci. U.S.A. 88, 4260-4264, 1991
A;Title: Isolation and properties of a nitrile hydratase from the soil fungus Myrotheciun
A;Reference number: A39365; MUID:91239547; PMID:2034671
A;Accession: A39365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-244 <MAI>
A;Cross-references: GB:M59078; NID:g168392; PIDN:AAA33429.1; PID:g168393
C;Superfamily: Saccharomyces cerevisiae hypothetical protein YFL061w
C;Keywords: carbon-oxygen lyase; hydro-lyase

  Query Match             80.0%;  Score 36;  DB 2;  Length 244;
  Best Local Similarity   66.7%;  Pred. No. 7.6;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        169 TLYDNVGAY 177


RESULT 7
S39296
hexon protein - human adenovirus 4
C;Species: Mastadenovirus h4 (human adenovirus 4)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S39296
R;Pring-Akerblom, P.; Adrian, T.
submitted to the EMBL Data Library, November 1993
A;Reference number: S39296
A;Accession: S39296
```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <PRI>
A;Cross-references: EMBL:X76550; NID:g434903; PIDN:CAA54052.1; PID:g434904
C;Superfamily: adenovirus hexon protein

Query Match 80.0%; Score 36; DB 2; Length 447;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|| |||:|
Db 355 FLYANVGLY 363

RESULT 8
F70190
probable diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90) - Lyme dise
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 04-Aug-2003
C;Accession: F70190
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: F70190
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-448 <KLE>
A;Cross-references: GB:AE001172; GB:AE000783; NID:g2688654; PIDN:AAC67070.1; PID:g268865
A;Experimental source: strain B31
C;Superfamily: pyrophosphate-dependent phosphofructokinase, Eh/PPi-PFK type; 6-phosphofr
C;Keywords: phosphotransferase
F;82-398/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 80.0%; Score 36; DB 1; Length 448;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|||::|:|
Db 336 LLYEDIGLY 344

RESULT 9
HMIV2
hemagglutinin precursor - influenza A virus (strain A/Japan/305/57[H2])
C;Species: influenza A virus
A;Variety: strain A/Japan/305/57[H2]
C;Date: 28-Feb-1981 #sequence_revision 28-Feb-1981 #text_change 16-Jul-1999
C;Accession: A04062; S12270
R;Gething, M.J.; Bye, J.; Skehel, J.; Waterfield, M.
Nature 287, 301-306, 1980
A;Title: Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from
A;Reference number: A93233; MUID:81030852; PMID:7421990
A;Accession: A04062
A;Molecule type: mRNA
A;Residues: 1-562 <GET>
A;Cross-references: GB:J02127; NID:g324145; PIDN:AAA43185.1; PID:g324146
A;Experimental source: strain A/Japan/305/57[H2]
R;Naeve, C.W.; Williams, D.
EMBO J. 9, 3857-3866, 1990
A;Title: Fatty acids on the A/Japan/305/57 influenza virus hemagglutinin have a role in
A;Reference number: S12270; MUID:91065313; PMID:2249653
A;Accession: S12270
A;Molecule type: mRNA
A;Residues: 510-562 <NAE>
A;Experimental source: strain A/Japan/305/57 (H2N2)
C;Superfamily: influenza virus hemagglutinin
C;Keywords: hemagglutinin; homotrimer; lipoprotein; thiolester bond
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-339/Product: hemagglutinin chain HA1 #status predicted <HA1>
F;341-562/Product: hemagglutinin chain HA2 #status predicted <HA2>
F;551,558,561/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 80.0%; Score 36; DB 1; Length 562;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:||:|||_|
Db 203 TLYQNVGTY 211

RESULT 10
S57637
hexon protein - human adenovirus 4
C;Species: Mastadenovirus h4 (human adenovirus 4)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C;Accession: S57637
R;Pring-Akerblom, P.; Trijssenaar, J.; Adrian, T.
submitted to the EMBL Data Library, February 1995
A;Reference number: S57637
A;Accession: S57637
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-936 <PRI>
A;Cross-references: EMBL:X84646; NID:g886486; PIDN:CAA59139.1; PID:g886487
C;Superfamily: adenovirus hexon protein

Query Match 80.0%; Score 36; DB 2; Length 936;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|| |||:|
Db 456 FLYANVGLY 464

RESULT 11
B81136
hypothetical protein NMB0968 [imported] - Neisseria meningitidis (strain MC58 serogroup B
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: B81136
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: B81136
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-29 <TET>
A;Cross-references: GB:AE002448; GB:AE002098; NID:g7226204; PIDN:AAF41373.1; PID:g722620
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0968

Query Match 77.8%; Score 35; DB 2; Length 29;
Best Local Similarity 55.6%; Pred. No. 1.1;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:||:|:|:|
Db 21 FLYKNLGLY 29

RESULT 12
H81883
hypothetical protein NMA1165 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: H81883
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: H81883
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-34 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84427.1; PID:g737985
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1165

Query Match 77.8%; Score 35; DB 2; Length 34;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:|||:|:|:|
Db 26 FLYKNLGLY 34

RESULT 13
A55883
actin-filament-associated protein 120k form - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C;Accession: A55883
R;Flynn, D.C.; Koay, T.C.; Humphries, C.G.; Guappone, A.C.
J. Biol. Chem. 270, 3894-3899, 1995
A;Title: AFAP-120. A variant form of the Src SH2/SH3-binding partner AFAP-110 is detecte
A;Reference number: A55883; MUID:95181352; PMID:7876134
A;Accession: A55883
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-150 <FLY>
A;Cross-references: GB:L20302

Query Match 77.8%; Score 35; DB 2; Length 150;
Best Local Similarity 55.6%; Pred. No. 7.1;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:||:| |:|
Db 51 MLYDNAGLY 59

RESULT 14
A99574
ABC transporter atp-binding protein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C;Accession: A99574
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: A99574
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <KUR>
A;Cross-references: GB:AL445566; PID:g14089911; PIDN:CAC13670.1; GSPDB:GN00153
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU_4970
A;Genetic code: SGC3

Query Match 75.6%; Score 34; DB 2; Length 511;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:||||: :|
Db 89 SLYENISVY 97

RESULT 15
A56511
myosin I myoA - Emericella nidulans
C;Species: Emericella nidulans, Aspergillus nidulans
C;Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 02-Feb-2001
C;Accession: A56511
R;McGoldrick, C.A.; Gruver, C.; May, G.S.
J. Cell Biol. 128, 577-587, 1995
A;Title: myoA of Aspergillus nidulans encodes an essential myosin I required for secretic
A;Reference number: A56511; MUID:95164560; PMID:7860631
A;Accession: A56511
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1249 <MCG>
A;Cross-references: GB:U12427; NID:g525321; PIDN:AAA67877.1; PID:g525322
C;Genetics:
A;Gene: myoA
C;Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homolog
C;Keywords: nucleotide binding; P-loop
F;53-716/Domain: myosin motor domain homology <MMOT>
F;143-150/Region: nucleotide-binding motif A (P-loop)
F;1081-1130/Domain: SH3 homology <SH3>

Query Match 75.6%; Score 34; DB 2; Length 1249;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMY 9
:||::||:|
Db 1013 DLYQSVGLY 1021

Search completed: July 15, 2004, 07:29:22
Job time : 13.5 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:20:47 ; Search time 8 Seconds
(without alignments)
58.579 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	86.7	565	1	PHR_YEAST	P05066 saccharomyc
2	37	82.2	468	1	HEX_ADE31	P36855 human adeno
3	37	82.2	526	1	VGLG_SIGMA	P12647 sigma virus
4	37	82.2	919	1	HEX_ADE12	P19900 human adeno
5	36	80.0	244	1	CYAH_MYRVE	P22143 myrothecium
6	36	80.0	447	1	HEX_ADE04	P36850 human adeno
7	36	80.0	562	1	HEMA_IAJAP	P03451 influenza a
8	35	77.8	1101	1	DIA2_HUMAN	O60879 homo sapien
9	33	73.3	99	1	YLW3_CAEEL	P34406 caenorhabdi
10	33	73.3	306	1	PYRB_METJA	Q58976 methanococc
11	33	73.3	312	1	FDXH_HAEIN	P44450 haemophilus
12	33	73.3	313	1	CEO2_LACLA	P15244 lactococcus
13	33	73.3	512	1	VENV_THOGV	P28977 thogoto vir
14	33	73.3	693	1	AGLU_SULSO	O59645 sulfolobus
15	33	73.3	754	1	RAD4_YEAST	P14736 saccharomyc
16	32	71.1	221	1	Y805_METJA	Q58215 methanococc
17	32	71.1	307	1	METF_STRLI	O54235 streptomyce
18	32	71.1	437	1	PAAK_ECOLI	P76085 escherichia
19	32	71.1	450	1	DCOR_CHICK	P27118 gallus gall
20	32	71.1	455	1	DCOR_CRIGR	P14019 cricetulus
21	32	71.1	456	1	DCO2_XENLA	Q9i8s4 xenopus lae
22	32	71.1	460	1	DCOR_XENLA	P27120 xenopus lae
23	32	71.1	461	1	DCOR_BOVIN	P27117 bos taurus
24	32	71.1	461	1	DCOR_HUMAN	P11926 homo sapien
25	32	71.1	461	1	DCOR_MOUSE	P00860 mus musculu
26	32	71.1	461	1	DCOR_MUSPA	P27119 mus pahari
27	32	71.1	461	1	DCOR_RAT	P09057 rattus norv
28	32	71.1	519	1	ALGG_PSEPK	Q88nc9 pseudomonas
29	32	71.1	536	1	ALGG_PSESM	Q887q3 pseudomonas
30	32	71.1	593	1	PTNB_CHICK	Q90687 gallus gall
31	32	71.1	593	1	PTNB_HUMAN	Q06124 homo sapien
32	32	71.1	593	1	PTNB_RAT	P41499 rattus norv
33	32	71.1	671	1	RIK1_HUMAN	Q13546 homo sapien
34	32	71.1	726	1	NU84_YEAST	P52891 saccharomyc
35	32	71.1	1018	1	CONT_HUMAN	Q12860 homo sapien
36	32	71.1	1020	1	CONT_MOUSE	P12960 mus musculu
37	32	71.1	1021	1	CONT_RAT	Q63198 rattus norv
38	32	71.1	1364	1	PUR2_DROPS	P16340 d trifuncti
39	32	71.1	1900	1	STT4_YEAST	P37297 saccharomyc
40	31	68.9	177	1	YP98_CAEEL	Q09246 caenorhabdi
41	31	68.9	178	1	YQAC_BACSU	P45900 bacillus su
42	31	68.9	201	1	LEUD_BUCBP	P59516 buchnera ap
43	31	68.9	251	1	Y116_MYCPN	P75519 mycoplasma
44	31	68.9	342	1	K6P1_STRCO	O08333 streptomyce
45	31	68.9	360	1	CYSP_HEMSP	P43156 hemerocalli

ALIGNMENTS

RESULT 1
PHR_YEAST
ID PHR_YEAST STANDARD; PRT; 565 AA.
AC P05066;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribodipyrimidine photolyase, mitochondrial precursor
DE (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme).
GN PHR1 OR YOR386W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067229; PubMed=3906569;
RA Sancar G.B.;
RT "Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of
RT the PHR1 photolyase to E. coli photolyase.";
RL Nucleic Acids Res. 13:8231-8246(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86083177; PubMed=3000886;
RA Yasui A., Langeveld S.A.;
RT "Homology between the photoreactivation genes of Saccharomyces
RT cerevisiae and Escherichia coli.";
RL Gene 36:349-355(1985).
RN [3]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP REVIEW.
RA Sancar G.B., Sancar A.;
RT "Structure and function of DNA photolyases.";
RL Trends Biochem. Sci. 12:259-261(1987).
CC -!- FUNCTION: This enzyme catalyzes the light-dependent monomerization
CC (300-600 nM) of cyclobutyl pyrimidine dimers (in cis-syn
CC configuration), which are formed between adjacent bases on the
CC same DNA strand, upon exposure to ultraviolet radiation.
CC -!- CATALYTIC ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine
CC residues (in DNA).
CC -!- COFACTOR: Contains 2 chromophores: a reduced flavin (FADH2) and a
CC 5,10-methenyltetrahydrofolate. Both chromophores are bound by non-
CC covalent interactions.
CC -!- SUBCELLULAR LOCATION: Nuclear and mitochondrial.
CC -!- MISCELLANEOUS: This protein belongs to the "short wavelength-type
CC photolyases" with an absorption maximum at about 380 nM.
CC -!- MISCELLANEOUS: There are only 150-300 molecules of photolyase per
CC yeast cell.
CC -!- SIMILARITY: Belongs to the DNA photolyase class-1 family.
CC ---

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X03183; CAA26944.1; -.
DR   EMBL; M11578; AAA34875.1; -.
DR   EMBL; Z75294; CAA99718.1; -.
DR   PIR; S67298; S67298.
DR   HSSP; P00914; 1DNP.
DR   GermOnline; 143974; -.
DR   SGD; S0005913; PHR1.
DR   InterPro; IPR002081; DNA_photolyase_1.
DR   InterPro; IPR006050; DNA_photolyase_N.
DR   InterPro; IPR005101; FAD_binding_7.
DR   InterPro; IPR006051; FAD_binding_N.
DR   Pfam; PF00875; DNA_photolyase; 1.
DR   Pfam; PF03441; FAD_binding_7; 1.
DR   PRINTS; PR00147; DNAPHOTLYASE.
DR   ProDom; PD004390; FAD_binding_N; 1.
DR   PROSITE; PS00394; DNA_PHOTOLYASES_1_1; 1.
DR   PROSITE; PS00691; DNA_PHOTOLYASES_1_2; 1.
KW   Lyase; Chromophore; Flavoprotein; FAD; DNA repair; DNA-binding;
KW   Nuclear protein; Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION.
FT   CHAIN         ?    565       DEOXYRIBODIPYRIMIDINE PHOTOLYASE.
FT   DNA_BIND    421    440       H-T-H MOTIF (POTENTIAL).
FT   CONFLICT     77     77       V -> A (IN REF. 2).
FT   CONFLICT    165    165       T -> S (IN REF. 2).
FT   CONFLICT    169    169       S -> T (IN REF. 2).
FT   CONFLICT    200    200       D -> S (IN REF. 2).
FT   CONFLICT    351    351       S -> R (IN REF. 2).
FT   CONFLICT    365    365       G -> E (IN REF. 2).
FT   CONFLICT    473    473       E -> K (IN REF. 2).
SQ   SEQUENCE   565 AA;  66274 MW;  CD4FC3DA6128B97C CRC64;

  Query Match             86.7%;  Score 39;  DB 1;  Length 565;
  Best Local Similarity   66.7%;  Pred. No. 3;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:|||:|
Db         86 RLYDNVGLY 94


RESULT 2
HEX_ADE31
ID   HEX_ADE31      STANDARD;      PRT;   468 AA.
AC   P36855;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Hexon protein (Late protein 2) (Fragment).
GN   PII.
OS   Human adenovirus type 31.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=10529;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VRL 15/62;
RX   MEDLINE=94294642; PubMed=8023012;
RA   Pring-Akerblom P., Adrian T.;
RT   "Type- and group-specific polymerase chain reaction for adenovirus
RT   detection.";
RL   Res. Virol. 145:25-35(1994).
CC   -!- FUNCTION: This protein is one of the structural proteins in the
CC       viral coat and is synthesized during late infection.
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X74661; CAA52725.1; -.
DR   PIR; S37217; S37217.
DR   HSSP; P03277; 1DHX.
DR   InterPro; IPR000736; Adeno_hexon.
DR   Pfam; PF01065; Adeno_hexon; 1.
DR   ProDom; PD002815; Adeno_hexon; 1.
KW   Coat protein; Hexon protein; Late protein.
FT   NON_TER       1      1
FT   NON_TER     468    468
SQ   SEQUENCE   468 AA;  52100 MW;  8727BFA49179CE68 CRC64;

  Query Match             82.2%;  Score 37;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 6.3;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        341 FLYSNVGLY 349


RESULT 3
VGLG_SIGMA
ID   VGLG_SIGMA     STANDARD;      PRT;   526 AA.
AC   P12647;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Spike glycoprotein precursor.
GN   G.
OS   Sigma virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Rhabdoviridae; unclassified Rhabdoviridae.
OX   NCBI_TaxID=11301;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88034947; PubMed=2822842;
RA   Teninges D., Bras-Herreng F.;
RT   "Rhabdovirus sigma, the hereditary CO2 sensitivity agent of
RT   Drosophila: nucleotide sequence of a cDNA clone encoding the
RT   glycoprotein.";
RL   J. Gen. Virol. 68:2625-2638(1987).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X06171; CAA29536.1; -.
DR   PIR; A27150; VGVNSG.
DR   FlyBase; FBgn0015809; Sigma-Virus\G.
DR   InterPro; IPR001903; Rhabd_glycop.
DR   Pfam; PF00974; Rhabdo_glycop; 1.
KW   Transmembrane; Envelope protein; Glycoprotein; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    526       SPIKE GLYCOPROTEIN.
FT   CARBOHYD     32     32       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   526 AA;  59010 MW;  335607C69249DD9D CRC64;

  Query Match             82.2%;  Score 37;  DB 1;  Length 526;
  Best Local Similarity   66.7%;  Pred. No. 7.1;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||::||||
```

```
Db        350 VLYQSVGMY 358


RESULT 4
HEX_ADE12
ID   HEX_ADE12      STANDARD;      PRT;   919 AA.
AC   P19900;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hexon protein (Late protein 2).
GN   PII.
OS   Human adenovirus type 12.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=28282;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94076430; PubMed=8254750;
RA   Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;
RT   "Nucleotide sequence of human adenovirus type 12 DNA: comparative
RT   functional analysis.";
RL   J. Virol. 68:379-389(1994).
RN   [2]
RP   SEQUENCE OF 888-919 FROM N.A.
RC   STRAIN=Pereira 1131;
RX   MEDLINE=88303354; PubMed=3043380;
RA   Weber J.M., Houde A.;
RT   "The primary structure of human adenovirus type 12 protease.";
RL   Nucleic Acids Res. 16:7195-7195(1988).
CC   -!- FUNCTION: This protein is one of the structural proteins in the
CC       viral coat and is synthesized during late infection.
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X73487; CAA51891.1; -.
DR   EMBL; X07655; CAA30501.1; -.
DR   EMBL; X07655; CAB37192.1; -.
DR   PIR; S01730; S01730.
DR   PIR; S33942; S33942.
DR   HSSP; P03277; 1DHX.
DR   InterPro; IPR000736; Adeno_hexon.
DR   Pfam; PF01065; Adeno_hexon; 1.
DR   Pfam; PF03678; Adeno_hexon_C; 1.
DR   ProDom; PD002815; Adeno_hexon; 1.
KW   Coat protein; Hexon protein; Late protein.
SQ   SEQUENCE   919 AA;  103039 MW;  B37167885A516288 CRC64;

  Query Match             82.2%;  Score 37;  DB 1;  Length 919;
  Best Local Similarity   66.7%;  Pred. No. 13;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        439 FLYSNVGLY 447


RESULT 5
CYAH_MYRVE
ID   CYAH_MYRVE     STANDARD;      PRT;   244 AA.
AC   P22143;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Cyanamide hydratase (EC 4.2.1.69) (Urea hydro-lyase).
GN   CAH.
OS   Myrothecium verrucaria.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Myrothecium.
OX   NCBI_TaxID=5532;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=DSM 2087;
RX   MEDLINE=91239547; PubMed=2034671;
RA   Maier-Greiner U.M., Obermaier-Skrobranek B.M.M., Estermaier L.M.,
RA   Kammerloher W., Freund C., Wuelfing C., Burkert U.I., Matern D.H.,
RA   Breuer M., Eulitz M., Kuefrevioglu O.I., Hartmann G.R.;
RT   "Isolation and properties of a nitrile hydratase from the soil fungus
RT   Myrothecium verrucaria that is highly specific for the fertilizer
RT   cyanamide and cloning of its gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:4260-4264(1991).
CC   -!- FUNCTION: When used as herbicide in agriculture, cyanamide can be
CC       transformed, after sowing, in soil fertilizing ammonia by the
CC       combined action of M.verrucaria cyanamide hydratase and urease.
CC   -!- CATALYTIC ACTIVITY: Urea = cyanamide + H(2)O.
CC   -!- COFACTOR: Zinc.
CC   -!- SUBUNIT: Homohexamer.
CC   -!- MISCELLANEOUS: This enzyme is highly specific for cyanamide.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M59078; AAA33429.1; -.
DR   PIR; A39365; A39365.
DR   InterPro; IPR006674; HD.
DR   InterPro; IPR003607; Met_phsphohydro.
DR   Pfam; PF01966; HD; 1.
DR   SMART; SM00471; HDc; 1.
KW   Lyase; Zinc.
SQ   SEQUENCE   244 AA;  26966 MW;  880FA11F30E31CE2 CRC64;

  Query Match             80.0%;  Score 36;  DB 1;  Length 244;
  Best Local Similarity   66.7%;  Pred. No. 5;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        169 TLYDNVGAY 177


RESULT 6
HEX_ADE04
ID   HEX_ADE04      STANDARD;      PRT;   447 AA.
AC   P36850;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hexon protein (Late protein 2) (Fragment).
GN   PII.
OS   Human adenovirus type 4.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=28280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate RJ-67;
RX   MEDLINE=95407102; PubMed=7676636;
RA   Pring-Akerblom P., Trijssenaar J., Adrian T.;
RT   "Sequence characterization and comparison of human adenovirus
RT   subgenus B and E hexons.";
RL   Virology 212:232-236(1995).
CC   -!- FUNCTION: This protein is one of the structural proteins in the
CC       viral coat and is synthesized during late infection.
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X76550; CAA54052.1; -.
DR   PIR; S39296; S39296.
DR   HSSP; P03277; 1DHX.
DR   InterPro; IPR000736; Adeno_hexon.
DR   Pfam; PF01065; Adeno_hexon; 1.
DR   ProDom; PD002815; Adeno_hexon; 1.
KW   Coat protein; Hexon protein; Late protein.
FT   NON_TER      1     1
FT   NON_TER    447   447
SQ   SEQUENCE   447 AA;  49553 MW;  A7AE1977F707BD4D CRC64;

  Query Match             80.0%;  Score 36;  DB 1;  Length 447;
  Best Local Similarity   66.7%;  Pred. No. 9.6;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        355 FLYANVGLY 363


RESULT 7
HEMA_IAJAP
ID   HEMA_IAJAP     STANDARD;      PRT;   562 AA.
AC   P03451;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE   Hemagglutinin HA2 chain].
GN   HA.
OS   Influenza A virus (strain A/Japan/305/57).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=11421;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81030852; PubMed=7421990;
RA   Gething M.-J., Bye J., Skehel J.J., Waterfield M.;
RT   "Cloning and DNA sequence of double-stranded copies of haemagglutinin
RT   genes from H2 and H3 strains elucidates antigenic shift and drift in
RT   human influenza virus.";
RL   Nature 287:301-306(1980).
CC   -!- FUNCTION: Hemagglutinin is responsible for attaching the virus to
CC       cell receptors and for initiating infection.
CC   -!- SUBUNIT: Homotrimer. Each of the monomers is formed by two chains
CC       (HA1 and HA2) linked by a disulfide bond.
CC   -!- SIMILARITY: Belongs to the influenza viruses hemagglutinin family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J02127; AAA43185.1; -.
DR   PIR; A04062; HMIV2.
DR   HSSP; P03437; 1HTM.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT   SIGNAL       1     15
FT   CHAIN       16    339       HEMAGGLUTININ HA1 CHAIN.
FT   CHAIN      341    562       HEMAGGLUTININ HA2 CHAIN.
FT   CARBOHYD    25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   179    179       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   180    180       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   300    300       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   494    494       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   553    553       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   562 AA;  63118 MW;  6B7FD0C038993630 CRC64;

  Query Match             80.0%;  Score 36;  DB 1;  Length 562;
  Best Local Similarity   66.7%;  Pred. No. 12;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        203 TLYQNVGTY 211


RESULT 8
DIA2_HUMAN
ID   DIA2_HUMAN     STANDARD;      PRT;  1101 AA.
AC   O60879; O60878; Q9UJL2;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Diaphanous protein homolog 2 (Diaphanous-related formin 2) (DRF2).
GN   DIAPH2 OR DIA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=98163437; PubMed=9497258;
RA   Bione S., Sala C., Manzini C., Arrigo G., Zuffardi O., Banfi S.,
RA   Borsani G., Jonveaux P., Philippe C., Zuccotti M., Ballabio A.,
RA   Toniolo D.;
RT   "A human homologue of the Drosophila melanogaster diaphanous gene is
RT   disrupted in a patient with premature ovarian failure: evidence for
RT   conserved function in oogenesis and implications for human
RT   sterility.";
RL   Am. J. Hum. Genet. 62:533-541(1998).
RN   [2]
RP   SEQUENCE OF 685-906 FROM N.A.
RA   Heath P.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May be involved in oogenesis.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=DIA-156;
CC         IsoId=O60879-1; Sequence=Displayed;
CC       Name=DIA-12C;
CC         IsoId=O60879-2; Sequence=VSP_001573;
CC   -!- TISSUE SPECIFICITY: Expressed in testis, ovary, small intestine,
CC       prostate, lung, liver, kidney, Leukocytes.
CC   -!- DEVELOPMENTAL STAGE: Expressed from E16 in ovary and testis and
CC       during P6-P16 during differentiation of ovarian follicles.
CC   -!- DOMAIN: DRFs are regulated by intramolecular GBD-DAD binding where
CC       Rho-GTP activates the DRFs by disrupting the GBD-DAD interaction
CC       (By similarity).
CC   -!- DISEASE: Defects in DIAPH2 are a cause of premature ovarian
CC       failure (POF) [MIM:311360].
CC   -!- SIMILARITY: Contains 1 GTPase-binding (GBD) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 3 (FH3) domain.
CC   -!- SIMILARITY: Contains 1 DRF autoregulatory (DAD) domain.
CC   -!- SIMILARITY: Belongs to the formin homology family. Diaphanous
CC       subfamily.
CC   ----------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Y15909; CAA75870.1; -.
DR   EMBL; Y15908; CAA75869.1; -.
DR   EMBL; AL031053; CAB39108.1; -.
DR   Genew; HGNC:2877; DIAPH2.
DR   MIM; 300108; -.
DR   MIM; 311360; -.
DR   GO; GO:0005102; F:receptor binding; TAS.
DR   GO; GO:0016288; P:cytokinesis; TAS.
DR   GO; GO:0007292; P:female gamete generation; TAS.
DR   InterPro; IPR003104; FH2.
DR   Pfam; PF02181; FH2; 1.
DR   SMART; SM00498; FH2; 1.
KW   Alternative splicing; Coiled coil; Repeat.
FT   DOMAIN       86    285       GBD.
FT   DOMAIN      184    482       FH3.
FT   DOMAIN      366    418       COILED COIL (POTENTIAL).
FT   DOMAIN      487    547       COILED COIL (POTENTIAL).
FT   DOMAIN      549    623       FH1 (PRO-RICH).
FT   DOMAIN      628   1071       FH2.
FT   DOMAIN      903   1053       COILED COIL (POTENTIAL).
FT   DOMAIN     1054   1068       DAD.
FT   DOMAIN     1072   1075       ARG/LYS-RICH (BASIC).
FT   DOMAIN      257    260       POLY-LEU.
FT   DOMAIN      543    546       POLY-SER.
FT   DOMAIN      562    572       POLY-PRO.
FT   DOMAIN      576    585       POLY-PRO.
FT   DOMAIN      591    597       POLY-PRO.
FT   DOMAIN      603    608       POLY-PRO.
FT   DOMAIN      613    616       POLY-PRO.
FT   DOMAIN     1038   1041       POLY-LYS.
FT   VARSPLIC   1081   1101       DNRRVPLERSRSRHNGAISSK -> VVNHPCATRANPRSAT
FT                                (in isoform DIA-12C).
FT                                /FTId=VSP_001573.
SQ   SEQUENCE   1101 AA;  125568 MW;  399F1C292D79188B CRC64;

  Query Match             77.8%;  Score 35;  DB 1;  Length 1101;
  Best Local Similarity   66.7%;  Pred. No. 40;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||||:| |
Db        970 KLYENLGEY 978


RESULT 9
YLW3_CAEEL
ID   YLW3_CAEEL     STANDARD;      PRT;    99 AA.
AC   P34406;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein F22B7.3 in chromosome III.
GN   F22B7.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L12018; AAA65463.1; -.
DR   PIR; S44632; S44632.
DR   WormPep; F22B7.3; CE00156.
KW   Hypothetical protein.
SQ   SEQUENCE   99 AA;  11665 MW;  78FC94DBD3C8B585 CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 99;
  Best Local Similarity   71.4%;  Pred. No. 8.1;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YENVGMY 9
              |||:||:
Db         21 YENLGMF 27


RESULT 10
PYRB_METJA
ID   PYRB_METJA     STANDARD;      PRT;   306 AA.
AC   Q58976;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Aspartate carbamoyltransferase (EC 2.1.3.2) (Aspartate
DE   transcarbamylase) (ATCase).
GN   PYRB OR MJ1581.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=20283607; PubMed=10748118;
RA   Hack E.S., Vorobyova T., Sakash J.B., West J.M., Macol C.P., Herve G.,
RA   Williams M.K., Kantrowitz E.R.;
RT   "Characterization of the aspartate transcarbamoylase from
RT   Methanococcus jannaschii.";
RL   J. Biol. Chem. 275:15820-15827(2000).
RN   [3]
RP   CRYSTALLIZATION, AND X-RAY CRYSTALLOGRAPHY.
```

```
RX   MEDLINE=20402716; PubMed=10944354;
RA   Vitali J., Vorobyova T., Webster G., Kantrowitz E.R.;
RT   "Crystallization and structure determination of the catalytic trimer
RT   of Methanococcus jannaschii aspartate transcarbamoylase.";
RL   Acta Crystallogr. D 56:1061-1063(2000).
CC   -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
CC       + N-carbamoyl-L-aspartate.
CC   -!- PATHWAY: Pyrimidine biosynthesis; second step.
CC   -!- SUBUNIT: HETERODODECAMER (2C3:3R2) OF SIX CATALYTIC PYRB CHAINS
CC       ORGANIZED AS TWO TRIMERS (C3), AND SIX REGULATORY PYRI CHAINS
CC       ORGANIZED AS THREE DIMERS (R2).
CC   -!- SIMILARITY: Belongs to the ATCase/OTCase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U67598; AAB99601.1; -.
DR   PIR; D64497; D64497.
DR   HSSP; P00479; 3CSU.
DR   TIGR; MJ1581; -.
DR   HAMAP; MF_00001; -; 1.
DR   InterPro; IPR006130; Asp/Orn_COtranf.
DR   InterPro; IPR002082; Asp_carbmitransf.
DR   InterPro; IPR006131; OTCace_O.
DR   InterPro; IPR006132; OTCace_P.
DR   Pfam; PF00185; OTCace; 1.
DR   Pfam; PF02729; OTCace_N; 1.
DR   PRINTS; PR00100; AOTCASE.
DR   TIGRFAMs; TIGR00670; asp_carb_tr; 1.
DR   PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
KW   Pyrimidine biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   306 AA;  35159 MW;  CBDC31FC450CEF6A CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 306;
  Best Local Similarity   66.7%;  Pred. No. 26;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|:||| ||
Db        174 SLFENVEMY 182


RESULT 11
FDXH_HAEIN
ID   FDXH_HAEIN     STANDARD;      PRT;   312 AA.
AC   P44450;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Formate dehydrogenase, iron-sulfur subunit (Formate dehydrogenase beta
DE   subunit) (FDH beta subunit).
GN   FDXH OR HI0007.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: ALLOWS TO USE FORMATE AS MAJOR ELECTRON DONOR DURING
CC       ANAEROBIC RESPIRATION. THE BETA CHAIN IS AN ELECTRON TRANSPER UNIT
CC       CONTAINING 4 CYSTEINE CLUSTERS INVOLVED IN THE FORMATION OF IRON-
CC       SULFUR CENTRES. ELECTRONS ARE TRANSFERRED FROM THE GAMMA CHAIN TO
CC       THE MOLYBDENUM COFACTOR OF THE ALPHA SUBUNIT (BY SIMILARITY).
CC   -!- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
CC       BY SUBUNITS ALPHA, BETA AND GAMMA.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: ORTHOLOG OF BOTH E.COLI FDNH AND FDOH.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32686; AAC21685.1; -.
DR   PIR; A64042; A64042.
DR   HSSP; P00193; 1DUR.
DR   TIGR; HI0007; -.
DR   InterPro; IPR001450; 4Fe4S_ferredoxin.
DR   InterPro; IPR006470; FDH_beta.
DR   Pfam; PF00037; fer4; 1.
DR   TIGRFAMs; TIGR01582; FDH-beta; 1.
DR   PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW   Electron transport; 4Fe-4S; Iron-sulfur; Transmembrane;
KW   Complete proteome.
FT   METAL        44     44       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL        47     47       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL        50     50       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL        54     54       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL       106    106       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       109    109       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       114    114       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       118    118       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       139    139       IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT   METAL       142    142       IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT   METAL       145    145       IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT   METAL       149    149       IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT   METAL       166    166       IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT   METAL       169    169       IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT   METAL       181    181       IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT   METAL       185    185       IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
SQ   SEQUENCE   312 AA;  34068 MW;  AA49DD3C17064866 CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 312;
  Best Local Similarity   71.4%;  Pred. No. 27;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 YENVGMY 9
              ||| |:|
Db        214 YENAGLY 220


RESULT 12
CEO2_LACLA
ID   CEO2_LACLA     STANDARD;      PRT;   313 AA.
AC   P15244;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   N(5)-(carboxyethyl)ornithine synthase (EC 1.5.1.24) (N(5)-(L-1-
DE   carboxyethyl)-L-ornithine:NADP(+) oxidoreductase) (CEOS).
GN   CEO.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
```

```
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-15.
RC   STRAIN=K1-23; TRANSPOSON=Tn5306;
RX   MEDLINE=95263576; PubMed=7744873;
RA   Donkersloot J.A., Thompson J.;
RT   "Cloning, expression, sequence analysis, and site-directed
RT   mutagenesis of the Tn5306-encoded N5-(carboxyethyl)ornithine synthase
RT   from Lactococcus lactis K1.";
RL   J. Biol. Chem. 270:12226-12234(1995).
RN   [2]
RP   SEQUENCE OF 1-37.
RC   STRAIN=K1;
RX   MEDLINE=89255467; PubMed=2498334;
RA   Thompson J.;
RT   "N5-(L-1-carboxyethyl)-L-ornithine:NADP+ oxidoreductase from
RT   Streptococcus lactis. Purification and partial characterization.";
RL   J. Biol. Chem. 264:9592-9601(1989).
RN   [3]
RP   SEQUENCE OF 256-263, AND CHARACTERIZATION.
RC   STRAIN=K1;
RX   MEDLINE=20014035; PubMed=10548058;
RA   Sackett D.L., Ruvinov S.B., Thompson J.;
RT   "N5-(L-1-carboxyethyl)-L-ornithine synthase: physical and spectral
RT   characterization of the enzyme and its unusual low pKa fluorescent
RT   tyrosine residues.";
RL   Protein Sci. 8:2121-2129(1999).
RN   [4]
RP   FOLDING STUDIES.
RC   STRAIN=K1;
RX   MEDLINE=99456521; PubMed=10525296;
RA   Ruvinov S.B., Thompson J., Sackett D.L., Ginsburg A.;
RT   "Tetrameric N(5)-(L-1-carboxyethyl)-L-ornithine synthase: guanidine.
RT   HCl-induced unfolding and a low temperature requirement for
RT   refolding.";
RL   Arch. Biochem. Biophys. 371:115-123(1999).
CC   -!- CATALYTIC ACTIVITY: N(5)-(L-1-carboxyethyl)-L-ornithine + NADP(+)
CC       + H(2)O = L-ornithine + pyruvate + NADPH.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- MASS SPECTROMETRY: MW=35.355; METHOD=MALDI.
CC   -!- MISCELLANEOUS: In the reverse direction L-lysine can act instead
CC       of L-ornithine, more slowly, yielding N(6)-(L-1-carboxyethyl)-L-
CC       lysine.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U23376; AAA86385.1; -.
DR   PIR; A57499; A57499.
DR   InterPro; IPR007698; AlaDh_PNT_C.
DR   InterPro; IPR007886; AlaDh_PNT_N.
DR   Pfam; PF01262; AlaDh_PNT_C; 1.
DR   Pfam; PF05222; AlaDh_PNT_N; 1.
KW   Oxidoreductase; NADP.
FT   NP_BIND     171    176       NADPH (POTENTIAL).
FT   MUTAGEN      15     15       R->K: LOSS OF ACTIVITY.
SQ   SEQUENCE   313 AA;  35323 MW;  B17FE0F477113C77 CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 313;
  Best Local Similarity   55.6%;  Pred. No. 27;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              ::||| | |
Db        261 PIYENAGKY 269


RESULT 13
VENV_THOGV
ID   VENV_THOGV     STANDARD;      PRT;   512 AA.
AC   P28977;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Envelope glycoprotein precursor (Surface glycoprotein 75).
GN   P4.
OS   Thogoto virus (isolate SiAr 126) (Tho).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Thogotovirus.
OX   NCBI_TaxID=126796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92124738; PubMed=1733105;
RA   Morse M.A., Marriott A.C., Nuttall P.A.;
RT   "The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like
RT   virus) is related to the baculovirus glycoprotein GP64.";
RL   Virology 186:640-646(1992).
CC   -!- FUNCTION: POSSIBLE ROLE IN ENDOCYTOTIC FUSION EVENTS DURING
CC       INFECTION.
CC   -!- SUBUNIT: Monomer (Probable).
CC   -!- SIMILARITY: TO DHORI VIRUS ENVELOPE GLYCOPROTEIN AND TO
CC       BACULOVIRUSES MAJOR ENVELOPE GLYCOPROTEIN (P64/P67).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M77280; AAA47918.1; -.
DR   PIR; A40821; VGIVTH.
DR   InterPro; IPR004955; Baculo_gp64.
DR   Pfam; PF03273; Baculo_gp64; 1.
KW   Glycoprotein; Transmembrane; Signal.
FT   SIGNAL       1     15       POTENTIAL.
FT   CHAIN       16    512       ENVELOPE GLYCOPROTEIN.
FT   TRANSMEM   479    502       POTENTIAL.
FT   CARBOHYD   185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   289    289       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   378    378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   416    416       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   512 AA;  57550 MW;  0398FC36284A0DF1 CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 512;
  Best Local Similarity   55.6%;  Pred. No. 45;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |:|:|
Db        484 LLYGNIGVY 492


RESULT 14
AGLU_SULSO
ID   AGLU_SULSO     STANDARD;      PRT;   693 AA.
AC   O59645;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Alpha-glucosidase (EC 3.2.1.20) (Maltase).
GN   MALA OR SSO3051 OR C23_036.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-20 AND 552-561.
RC   STRAIN=98/2;
RX   MEDLINE=98155158; PubMed=9495770;
```

```
RA   Rolfsmeier M., Haseltine C., Bini E., Clark A., Blum P.;
RT   "Molecular characterization of the alpha-glucosidase gene (malA) from
RT   the hyperthermophilic archaeon Sulfolobus solfataricus.";
RL   J. Bacteriol. 180:1287-1295(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC   -!- FUNCTION: MAJOR SOLUBLE ALPHA-GLUCOSIDASE.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
CC       linked D-glucose residues with release of D-glucose.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- INDUCTION: EXPRESSED DURING GROWTH ON MALTOSE.
CC   -!- MISCELLANEOUS: THE PH OPTIMUM FOR MALTOSE HYDROLYSIS IS 4.5, AND
CC       5.5 FOR GLUCOGEN HYDROLYSIS.
CC   -!- SIMILARITY: Belongs to family 31 of glycosyl hydrolases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF042494; AAC38215.1; -.
DR   EMBL; AE006896; AAK43151.1; -.
DR   PIR; H90486; H90486.
DR   InterPro; IPR000322; Glyco_hydro_31.
DR   Pfam; PF01055; Glyco_hydro_31; 1.
DR   PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 1.
DR   PROSITE; PS00707; GLYCOSYL_HYDROL_F31_2; FALSE_NEG.
KW   Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    320    320       BY SIMILARITY.
SQ   SEQUENCE   693 AA;  80441 MW;  27BB952C0A7B3858 CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 693;
  Best Local Similarity   55.6%;  Pred. No. 62;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              ::||| |:|
Db          5 KIYENKGVY 13


RESULT 15
RAD4_YEAST
ID   RAD4_YEAST     STANDARD;      PRT;   754 AA.
AC   P14736;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   DNA repair protein RAD4.
GN   RAD4 OR YER162C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89232744; PubMed=3073107;
RA   Gietz R.D., Prakash S.;
RT   "Cloning and nucleotide sequence analysis of the Saccharomyces
RT   cerevisiae RAD4 gene required for excision repair of UV-damaged
RT   DNA.";
RL   Gene 74:535-541(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89197751; PubMed=2649477;
RA   Couto L.B., Friedberg E.C.;
RT   "Nucleotide sequence of the wild-type RAD4 gene of Saccharomyces
RT   cerevisiae and characterization of mutant rad4 alleles.";
RL   J. Bacteriol. 171:1862-1869(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313264; PubMed=9169868;
RA   Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA   Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA   Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA   Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA   Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA   Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA   Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL   Nature 387:78-81(1997).
CC   -!- FUNCTION: Involved in nucleotide excision repair of DNA damaged
CC       with UV light, bulky adducts, or cross-linking agents.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the XPC family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M26050; AAA34944.1; -.
DR   EMBL; M24928; AAA34945.1; -.
DR   EMBL; U18917; AAB64689.1; -.
DR   PIR; S30814; DDBYD4.
DR   GermOnline; 139239; -.
DR   SGD; S0000964; RAD4.
DR   GO; GO:0000111; C:nucleotide excision repair factor 2 complex; IDA.
DR   GO; GO:0000108; C:repairosome; IDA.
DR   GO; GO:0003684; F:damaged DNA binding; IDA.
DR   InterPro; IPR004583; Rad4.
DR   Pfam; PF03835; Rad4; 1.
DR   TIGRFAMs; TIGR00605; rad4; 1.
KW   DNA repair; DNA-binding; Nuclear protein.
FT   DNA_BIND    250    269       POTENTIAL.
FT   CONFLICT    223    225       VGI -> EGL (IN REF. 3).
SQ   SEQUENCE   754 AA;  87174 MW;  788C146DC4BD2BF8 CRC64;

  Query Match             73.3%;  Score 33;  DB 1;  Length 754;
  Best Local Similarity   71.4%;  Pred. No. 68;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YENVGMY 9
              |:|||:|
Db        220 YDNVGIY 226


Search completed: July 15, 2004, 07:27:01
Job time : 10 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:25:27 ; Search time 33 Seconds
(without alignments)
86.050 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMY 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	82.2	540	12	Q88452	Q88452 sigma virus
2	37	82.2	914	12	Q9IF30	Q9if30 bovine aden
3	36	80.0	339	12	Q9IFF9	Q9iff9 influenza a
4	36	80.0	339	12	Q9IFF4	Q9iff4 influenza a
5	36	80.0	339	12	Q9IFF2	Q9iff2 influenza a
6	36	80.0	339	12	Q9IFF6	Q9iff6 influenza a
7	36	80.0	339	12	Q9IFF0	Q9iff0 influenza a
8	36	80.0	339	12	Q9IFF7	Q9iff7 influenza a
9	36	80.0	339	12	Q9IFF8	Q9iff8 influenza a
10	36	80.0	339	12	Q9IFF5	Q9iff5 influenza a
11	36	80.0	339	12	Q9IFG0	Q9ifg0 influenza a
12	36	80.0	339	12	Q9IFF3	Q9iff3 influenza a
13	36	80.0	339	12	Q9IFF1	Q9iff1 influenza a
14	36	80.0	339	12	Q9IFG2	Q9ifg2 influenza a
15	36	80.0	339	12	Q9IFG1	Q9ifg1 influenza a
16	36	80.0	359	12	Q997B2	Q997b2 influenza a
17	36	80.0	359	12	Q997B3	Q997b3 influenza a
18	36	80.0	359	12	Q997B4	Q997b4 influenza a
19	36	80.0	359	12	Q997B1	Q997b1 influenza a
20	36	80.0	373	12	Q9WQX2	Q9wqx2 influenza a
21	36	80.0	376	12	Q9WQX1	Q9wqx1 influenza a
22	36	80.0	376	12	Q9WQW1	Q9wqw1 influenza a
23	36	80.0	376	12	Q9WQW4	Q9wqw4 influenza a
24	36	80.0	378	12	Q9WQX0	Q9wqx0 influenza a
25	36	80.0	378	12	Q9WQW8	Q9wqw8 influenza a
26	36	80.0	378	12	Q9WQW6	Q9wqw6 influenza a
27	36	80.0	378	12	Q9WQW2	Q9wqw2 influenza a
28	36	80.0	379	12	Q9WQX3	Q9wqx3 influenza a
29	36	80.0	380	12	Q9WQV9	Q9wqv9 influenza a
30	36	80.0	381	12	Q9WQW7	Q9wqw7 influenza a
31	36	80.0	381	12	Q9WQW5	Q9wqw5 influenza a
32	36	80.0	381	12	Q9WQW3	Q9wqw3 influenza a
33	36	80.0	381	12	Q9WQW0	Q9wqw0 influenza a
34	36	80.0	448	16	O51669	O51669 borrelia bu
35	36	80.0	560	12	Q9WQW9	Q9wqw9 influenza a
36	36	80.0	562	12	Q67032	Q67032 influenza a
37	36	80.0	562	12	Q67085	Q67085 influenzavi
38	36	80.0	562	12	Q67208	Q67208 influenza a
39	36	80.0	562	12	Q67120	Q67120 influenzavi
40	36	80.0	562	12	Q67011	Q67011 influenza a
41	36	80.0	562	12	Q67284	Q67284 influenza a
42	36	80.0	562	12	Q67165	Q67165 influenza a
43	36	80.0	562	12	Q67143	Q67143 influenza a
44	36	80.0	562	12	Q67140	Q67140 influenza a
45	36	80.0	562	12	Q67326	Q67326 influenza a

ALIGNMENTS

RESULT 1
Q88452
ID Q88452 PRELIMINARY; PRT; 540 AA.
AC Q88452;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glycoprotein.
GN G.
OS Sigma virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; unclassified Rhabdoviridae.
OX NCBI_TaxID=11301;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=234HRC;
RX MEDLINE=96074506; PubMed=7491755;
RA Landes-Devauchelle C., Bras F., Dezelee S., Teninges D.;
RT "Gene 2 of the sigma rhabdovirus genome encodes the P protein, and
RT gene 3 encodes a protein related to the reverse transcriptase of
RT retroelements.";
RL Virology 213:300-312(1995).
DR EMBL; X91062; CAA62517.1; -.
DR InterPro; IPR001903; Rhabd_glycop.
DR Pfam; PF00974; Rhabdo_glycop; 1.
SQ SEQUENCE 540 AA; 60771 MW; 7A0B553D1EA5E98A CRC64;

Query Match 82.2%; Score 37; DB 12; Length 540;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMY 9
              :||::||||
Db        364 VLYQSVGMY 372
```

RESULT 2
Q9IF30
ID Q9IF30 PRELIMINARY; PRT; 914 AA.

```
AC   Q9IF30;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hexon protein (Fragment).
OS   Bovine adenovirus type 10 (Mastadenovirus bos10).
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=39788;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=78-5371;
RA   Lehmkuhl H.D., Hobbs L.A.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF282774; AAF82136.1; -.
DR   HSSP; P03277; 1DHX.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000736; Adeno_hexon.
DR   Pfam; PF01065; Adeno_hexon; 1.
DR   Pfam; PF03678; Adeno_hexon_C; 1.
DR   ProDom; PD002815; Adeno_hexon; 1.
FT   NON_TER     914    914
SQ   SEQUENCE   914 AA;  103905 MW;  5508E006997739CD CRC64;

  Query Match             82.2%;  Score 37;  DB 12;  Length 914;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :|| |||:|
Db        433 FLYSNVGLY 441


RESULT 3
Q9IFF9
ID   Q9IFF9      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Davis/1/57(H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=220951;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Davis/1/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270719; AAF82103.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37810 MW;  7D35925ED7538B08 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 4
Q9IFF4
ID   Q9IFF4      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Malaya/16/58(H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=220954;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Malaya/16/58;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270724; AAF82108.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37893 MW;  D59A261E1EB9B621 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 5
Q9IFF2
ID   Q9IFF2      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Victoria/15681/59(H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=220956;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Victoria/15681/59;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
```

```
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270726; AAF82110.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37964 MW;  97239D60CD1FFD08 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 6
Q9IFF6
ID   Q9IFF6      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/RI/5+/57 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135328;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/RI/5+/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270722; AAF82106.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37853 MW;  7C70576EBB5B2FC0 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 7
Q9IFF0
ID   Q9IFF0      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Chile/6/57 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135323;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Chile/6/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270728; AAF82112.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37810 MW;  7D35925ED7538B08 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 8
Q9IFF7
ID   Q9IFF7      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (strain A/Ann Arbor/6/60).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135322;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Ann Arbor/6/60;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270721; AAF82105.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
```

```
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37896 MW;  FECE7718D2628F0E CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 9
Q9IFF8
ID   Q9IFF8       PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Albany/7/57 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135321;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Albany/7/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270720; AAF82104.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37825 MW;  0D3E767F9241AA30 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 10
Q9IFF5
ID   Q9IFF5       PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Albany/6/58(H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=220953;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Albany/6/58;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270723; AAF82107.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37753 MW;  2ADC4BA8C590ADCE CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 11
Q9IFG0
ID   Q9IFG0       PRELIMINARY;      PRT;   339 AA.
AC   Q9IFG0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/RI/5-/57 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/RI/5-/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270718; AAF82102.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37798 MW;  FE7698C4DC1D15E6 CRC64;
```

```
  Query Match            80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 12
Q9IFF3
ID   Q9IFF3      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Sao Paolo/3/59 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135330;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Sao Paolo/3/59;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270725; AAF82109.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37895 MW;  97D69D60CD5AFD08 CRC64;

  Query Match            80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 13
Q9IFF1
ID   Q9IFF1      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFF1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Ohio/2/59 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135327;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Ohio/2/59;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270727; AAF82111.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37991 MW;  F6BC8A0403FD40CC CRC64;

  Query Match            80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


RESULT 14
Q9IFG2
ID   Q9IFG2      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFG2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/El Salvador/2/57 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=135325;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/El Salvador/2/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270716; AAF82100.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37874 MW;  237050D99292320A CRC64;

  Query Match            80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211
```

```
RESULT 15
Q9IFG1
ID   Q9IFG1      PRELIMINARY;      PRT;   339 AA.
AC   Q9IFG1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemagglutinin (Fragment).
OS   Influenza A virus (A/Leningrad/134/57 (H2N2)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza A viruses; Influenzavirus A.
OX   NCBI_TaxID=128148;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/Leningrad/134/57;
RA   Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA   Castrucci M.R., Donatelli I., Kawaoka Y.;
RT   "Early alterations of the receptor-binding properties of H1, H2 and H3
RT   avian influenza virus hemagglutinins after their introduction into
RT   mammals.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC       CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC       (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR   EMBL; AF270717; AAF82101.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR008980; Capsid_hemag.
DR   InterPro; IPR001364; Hemagglutn.
DR   Pfam; PF00509; Hemagglutinin; 1.
DR   PRINTS; PR00329; HEMAGGLUTN12.
DR   ProDom; PD000225; Hemagglutn; 1.
KW   Envelope protein; Glycoprotein; Hemagglutinin.
FT   NON_TER     339    339
SQ   SEQUENCE   339 AA;  37825 MW;  EC97187675C23218 CRC64;

  Query Match             80.0%;  Score 36;  DB 12;  Length 339;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMY 9
              :||:||| |
Db        203 TLYQNVGTY 211


Search completed: July 15, 2004, 07:30:40
Job time : 35 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:21:32 ; Search time 48 Seconds
(without alignments)
52.978 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	100.0	9	4	AAB48919	Aab48919 Generic S
2	39	100.0	9	4	AAB48917	Aab48917 SH2 domai
3	39	100.0	9	4	AAB48922	Aab48922 SH2 domai
4	39	100.0	9	5	ABG68582	Abg68582 Peptide G
5	39	100.0	10	4	AAB48923	Aab48923 SH2 domai
6	39	100.0	10	4	AAB48920	Aab48920 SH2 domai
7	39	100.0	10	4	AAB48926	Aab48926 SH2 domai
8	39	100.0	10	4	AAB48921	Aab48921 SH2 domai
9	39	100.0	10	4	AAB48928	Aab48928 SH2 domai
10	39	100.0	11	2	AAW46897	Aaw46897 G1C-S pep
11	39	100.0	11	2	AAW46896	Aaw46896 Non-phosp
12	39	100.0	11	5	ABG68419	Abg68419 G1 peptid
13	39	100.0	11	5	ABG68583	Abg68583 Peptide G
14	39	100.0	26	4	AAB48932	Aab48932 SH2 domai
15	39	100.0	26	4	AAB48933	Aab48933 SH2 domai
16	33	84.6	10	4	AAB48925	Aab48925 SH2 domai
17	33	84.6	10	4	AAB48927	Aab48927 SH2 domai
18	33	84.6	475	6	ABU40815	Abu40815 Protein e
19	33	84.6	586	7	ADC39248	Adc39248 Novel hum
20	33	84.6	592	7	ADE86367	Ade86367 Human PTP
21	33	84.6	593	2	AAR52991	Aar52991 Human pro
22	33	84.6	593	2	AAR99313	Aar99313 Human SH-
23	33	84.6	593	2	AAY13476	Aay13476 Peptide S
24	33	84.6	593	4	AAB59218	Aab59218 SHP-2 act
25	33	84.6	593	4	AAB59223	Aab59223 SHP-2 act
26	33	84.6	593	4	AAB59226	Aab59226 SHP-2 act
27	33	84.6	593	4	AAB59220	Aab59220 SHP-2 act
28	33	84.6	593	4	AAB59225	Aab59225 SHP-2 act
29	33	84.6	593	4	AAB59222	Aab59222 SHP-2 act
30	33	84.6	593	4	AAB59227	Aab59227 SHP-2 act
31	33	84.6	593	4	AAB59213	Aab59213 SHP-2 mut
32	33	84.6	593	4	AAB59215	Aab59215 SHP-2 pro
33	33	84.6	593	4	AAB59224	Aab59224 SHP-2 act
34	33	84.6	593	4	AAB59221	Aab59221 SHP-2 act
35	33	84.6	593	4	AAB59214	Aab59214 SHP-2 mut
36	33	84.6	593	4	AAB59219	Aab59219 SHP-2 act
37	33	84.6	593	7	ADE86321	Ade86321 Human PTP
38	33	84.6	593	7	ADE86372	Ade86372 Human PTP
39	33	84.6	593	7	ADE86374	Ade86374 Human PTP
40	33	84.6	593	7	ADE86355	Ade86355 Human PTP
41	33	84.6	593	7	ADE86358	Ade86358 Human PTP
42	33	84.6	593	7	ADE86369	Ade86369 Human PTP
43	33	84.6	593	7	ADE86364	Ade86364 Human PTP
44	33	84.6	593	7	ADE86363	Ade86363 Human PTP
45	33	84.6	593	7	ADE86371	Ade86371 Human PTP

ALIGNMENTS

RESULT 1
AAB48919
ID AAB48919 standard; peptide; 9 AA.
XX
AC AAB48919;
XX
DT 16-MAR-2001 (first entry)
XX
DE Generic SH2 domain cyclic peptide inhibitor, SEQ ID NO:3.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .9
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety, thereby
FT cyclising the peptide"
FT Misc-difference 1
FT /note= "Any naturally or non-naturally occurring amino
FT acid except Glu"
FT Modified-site 9
FT /note= "C-terminal amide"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
DR WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Disclosure; Page 5; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding

```
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence is a generic
CC   representation of a cyclic peptide of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.4e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 2
AAB48917
ID   AAB48917 standard; peptide; 9 AA.
XX
AC   AAB48917;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain cyclic peptide inhibitor, SEQ ID NO:1.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .9
FT                   /note= "The nitrogen atoms of the N-terminus and the C-
FT                   terminal amide are joined via a bridging moiety, thereby
FT                   cyclising the peptide"
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
FT   Modified-site   9
FT                   /note= "C-terminal amide"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Claim 1; Page 21; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a cyclic
CC   peptide of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.4e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 3
AAB48922
ID   AAB48922 standard; peptide; 9 AA.
XX
AC   AAB48922;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:7.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid; the nitrogen atom of
FT                   the N-terminus is joined to a ClCH2C(O) moiety"
FT   Modified-site   9
FT                   /note= "The carbon atom of the C-terminus is joined to a
FT                   C(CH2SH)C(O)NH2 moiety"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
```

PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Example 1; Page 13; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 39;  DB 4;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.4e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9
```

RESULT 4
ABG68582
ID ABG68582 standard; peptide; 9 AA.
XX
AC ABG68582;
XX
DT 07-OCT-2002 (first entry)
XX
DE Peptide G1TE #1.
XX
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; G1;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
PN WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI Krag DN, Pero SC, Oligino L;
XX
DR WPI; 2002-547451/58.
XX
PT Treatment or prophylaxis of a subject having a disorder characterized by
PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Fig 9B; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal
CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC placenta, colon, ovary, testes and lung. The present sequence is a G1
CC peptide (not defined) or derivative which is used to illustrate the
CC possible structures of cyclic Grb7 antagonists
XX
SQ Sequence 9 AA;

```
  Query Match             100.0%;  Score 39;  DB 5;  Length 9;
  Best Local Similarity   77.8%;  Pred. No. 1.4e+06;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          1 ELYENVGMY 9
```

RESULT 5
AAB48923
ID AAB48923 standard; peptide; 10 AA.
XX
AC AAB48923;
XX
DT 16-MAR-2001 (first entry)
XX
DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:8.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytostatic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .10
FT /note= "The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety, thereby
FT cyclising the peptide"
FT Modified-site 1
FT /label= Aad
FT Modified-site 10
FT /note= "C-terminal amide"
XX
PN WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX

```
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 2; Page 13; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a cyclic
CC   peptide of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.09;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 6
AAB48920
ID   AAB48920 standard; peptide; 10 AA.
XX
AC   AAB48920;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain cyclic peptide inhibitor, SEQ ID NO:4.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .10
FT                   /note= "The nitrogen atoms of the N-terminus and the C-
FT                   terminal amide are joined via a bridging moiety C(O)-CH2-
FT                   S-CH2-CHC(O)NH2, thereby cyclising the peptide"
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
FT   Modified-site   10
FT                   /note= "C-terminal amide"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 1; Page 12; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a cyclic
CC   peptide of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.09;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 7
AAB48926
ID   AAB48926 standard; peptide; 10 AA.
XX
AC   AAB48926;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:11.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   10
FT                   /label= Nle
FT                   /note= "C-terminal amide, joined to a solid matrix"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
```

```
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 4; Page 14; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a linear
CC   precursor of a peptide of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 10;
  Best Local Similarity   77.8%;  Pred. No. 0.09;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          1 ELYENVGMY 9


RESULT 8
AAB48921
ID   AAB48921 standard; peptide; 10 AA.
XX
AC   AAB48921;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:5.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 1; Page 12; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a linear
CC   precursor of a peptide of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.09;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 9
AAB48928
ID   AAB48928 standard; peptide; 10 AA.
XX
AC   AAB48928;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:14.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   10
FT                   /label= Aad
FT                   /note= "C-terminal amide, joined to a solid matrix"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
```

```
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 5; Page 15; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a linear
CC   precursor of a peptide of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 10;
  Best Local Similarity   77.8%;  Pred. No. 0.09;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          1 ELYENVGMY 9


RESULT 10
AAW46897
ID   AAW46897 standard; peptide; 11 AA.
XX
AC   AAW46897;
XX
DT   19-JUN-1998  (first entry)
XX
DE   G1C-S peptide.
XX
KW   SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
KW   signal transduction protein; non-phosphorylated; inhibition; treatment;
KW   hyper-proliferative disease; human cancer.
XX
OS   Unidentified.
XX
PN   WO9802176-A1.
XX
PD   22-JAN-1998.
XX
PF   16-JUL-1997;   97WO-US012501.
XX
PR   16-JUL-1996;   96US-0021858P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
PA   (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI   King CR,  Sastry L,  Krag D,  Oligino L;
XX
DR   WPI; 1998-110340/10.
XX
PT   Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
PT   transducing protein - at least as well as natural phosphorylated target,
PT   particularly from treatment of cancer.
XX
PS   Disclosure; Page 18; 39pp; English.
XX
CC   The present sequence represents a peptide designated G1C-S. This peptide
CC   is essentially the same as a non-phosphorylated peptide, G1, that is
CC   capable of binding to the src homology 2 (SH2) domain of Grb2, except
CC   that the terminal Cys residues of G1 are replaced with Ser residues. Grb2
CC   is a signal transduction protein. The binding affinity of the present
CC   peptide with Grb2 was tested, and it was demonstrated that the disulphide
CC   bond of G1 may be important. The G1 peptide binds to the SH2 domain of
CC   Grb2 with affinity similar to, or greater than, that of a SHC
CC   phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
CC   that has not been modified by phosphate or similar charged group. The G1
CC   peptide is used to inhibit a signal transduction process that involves
CC   binding of a phosphorylated protein or peptide to the SH2 domain of a
CC   signal transduction protein, particularly Grb2. It is used specifically
CC   for treatment of hyper-proliferative diseases, especially human cancer
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 39;  DB 2;  Length 11;
  Best Local Similarity   77.8%;  Pred. No. 0.1;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          2 ELYENVGMY 10


RESULT 11
AAW46896
ID   AAW46896 standard; peptide; 11 AA.
XX
AC   AAW46896;
XX
DT   19-JUN-1998  (first entry)
XX
DE   Non-phosphorylated peptide which binds to the SH2 domain of Grb2.
XX
KW   SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
KW   signal transduction protein; non-phosphorylated; inhibition; treatment;
KW   hyper-proliferative disease; human cancer; cyclic.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  1. .11
XX
PN   WO9802176-A1.
XX
PD   22-JAN-1998.
XX
PF   16-JUL-1997;   97WO-US012501.
XX
PR   16-JUL-1996;   96US-0021858P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
PA   (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI   King CR,  Sastry L,  Krag D,  Oligino L;
XX
DR   WPI; 1998-110340/10.
XX
PT   Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
PT   transducing protein - at least as well as natural phosphorylated target,
PT   particularly from treatment of cancer.
XX
```

PS Claim 9; Page 17; 39pp; English.
XX
CC The present sequence represents non-phosphorylated peptide, G1, that is
CC capable of binding to the src homology 2 (SH2) domain of Grb2. Grb2 is a
CC signal transduction protein. The G1 peptide binds to the SH2 domain of
CC Grb2 with affinity similar to, or greater than, that of a SHC
CC phosphopeptide (AAW46895). The G1 peptide contains a tyrosine residue
CC that has not been modified by phosphate or similar charged group. The G1
CC peptide is used to inhibit a signal transduction process that involves
CC binding of a phosphorylated protein or peptide to the SH2 domain of a
CC signal transduction protein, particularly Grb2. It is used specifically
CC for treatment of hyper-proliferative diseases, especially human cancer
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 39; DB 2; Length 11;
Best Local Similarity 77.8%; Pred. No. 0.1;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
:|||||||:
Db 2 ELYENVGMY 10

RESULT 12
ABG68419
ID ABG68419 standard; peptide; 11 AA.
XX
AC ABG68419;
XX
DT 07-OCT-2002 (first entry)
XX
DE G1 peptide.
XX
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; G1;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
PN WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI Krag DN, Pero SC, Oligino L;
XX
DR WPI; 2002-547451/58.
XX
PT Treatment or prophylaxis of a subject having a disorder characterized by
PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Page 102; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal
CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC placenta, colon, ovary, testes and lung. The present sequence is a G1
CC peptide (not defined) or derivative which is used to illustrate the
CC possible structures of cyclic Grb7 antagonists
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 39; DB 5; Length 11;
Best Local Similarity 77.8%; Pred. No. 0.1;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
:|||||||:
Db 2 ELYENVGMY 10

RESULT 13
ABG68583
ID ABG68583 standard; peptide; 11 AA.
XX
AC ABG68583;
XX
DT 07-OCT-2002 (first entry)
XX
DE Peptide G1TE #2.
XX
KW Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytostatic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; G1;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
PN WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PI Krag DN, Pero SC, Oligino L;
XX
DR WPI; 2002-547451/58.
XX
PT Treatment or prophylaxis of a subject having a disorder characterized by
PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Fig 9C; 186pp; English.
XX
CC The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phage display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal

```
CC   interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC   cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC   gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC   placenta, colon, ovary, testes and lung. The present sequence is a G1
CC   peptide (not defined) or derivative which is used to illustrate the
CC   possible structures of cyclic Grb7 antagonists
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 39;  DB 5;  Length 11;
  Best Local Similarity   77.8%;  Pred. No. 0.1;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          2 ELYENVGMY 10


RESULT 14
AAB48932
ID   AAB48932 standard; peptide; 26 AA.
XX
AC   AAB48932;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain peptide inhibitor linear precursor, SEQ ID NO:18.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; linear precursor.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 12; Page 19; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
CC   the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC   of less than 4.0 micromolar when the target protein is Grb2 (growth
CC   factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC   turn conformation. The peptides, and compositions comprising the
CC   peptides, are useful for inhibiting the binding of the SH2 domain to a
CC   target protein. They are particularly useful for preventing cancer,
CC   especially breast cancer. The present sequence represents a linear
CC   precursor of a peptide of the invention
XX
SQ   Sequence 26 AA;

  Query Match             100.0%;  Score 39;  DB 4;  Length 26;
  Best Local Similarity   88.9%;  Pred. No. 0.28;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 15
AAB48933
ID   AAB48933 standard; peptide; 26 AA.
XX
AC   AAB48933;
XX
DT   16-MAR-2001  (first entry)
XX
DE   SH2 domain cyclic peptide inhibitor, SEQ ID NO:19.
XX
KW   SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW   cytostatic; tumour; breast cancer; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .10
FT                   /note= "The nitrogen atom of the N-terminus and the Cys
FT                   10 sidechain are joined via a bridging moiety, thereby
FT                   cyclising part of the peptide"
FT   Modified-site   1
FT                   /note= "Gamma-carboxyglutamic acid"
XX
PN   WO200073326-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015201.
XX
PR   02-JUN-1999;   99US-0137187P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Roller PP,  Long Y,  Lung FT,  King CR,  Yang D;
XX
DR   WPI; 2001-137633/14.
XX
PT   Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT   homology 2 domain binding to target protein, useful for preventing
PT   cancer, especially breast cancer.
XX
PS   Example 12; Page 20; 26pp; English.
XX
CC   The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC   which bind to Src homology 2 (SH2) domains, preventing them from binding
CC   to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC   cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC   -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
CC   Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC   aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
CC   is either Aad or Glu. Optionally, there is a conservative or neutral
CC   amino acid substitution at either or both of Leu2 and Gly7, and
CC   optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC   The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC   -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC   which links the nitrogen atom of the N terminus to the nitrogen atom of
```

CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides, and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a cyclic
CC peptide of the invention
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 39; DB 4; Length 26;
Best Local Similarity 88.9%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9
```

Search completed: July 15, 2004, 07:28:50
Job time : 49 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:26:37 ; Search time 14.5 Seconds
(without alignments)
32.044 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	84.6	485	4	US-09-543-681A-4935	Sequence 4935, Ap
2	33	84.6	566	2	US-08-272-255-8	Sequence 8, Appli
3	33	84.6	566	5	PCT-US95-08565-8	Sequence 8, Appli
4	33	84.6	593	1	US-08-202-389-12	Sequence 12, Appl
5	33	84.6	593	1	US-08-018-129-5	Sequence 5, Appli
6	33	84.6	593	2	US-08-448-250-5	Sequence 5, Appli
7	33	84.6	593	4	US-09-282-257-5	Sequence 5, Appli
8	32	82.1	214	4	US-09-489-039A-12637	Sequence 12637, A
9	31	79.5	325	4	US-09-134-001C-3551	Sequence 3551, Ap
10	31	79.5	335	4	US-09-134-000C-3814	Sequence 3814, Ap
11	31	79.5	919	2	US-08-788-674-4	Sequence 4, Appli
12	30	76.9	19	4	US-09-376-343-3	Sequence 3, Appli
13	30	76.9	20	2	US-08-480-190-38	Sequence 38, Appl
14	30	76.9	20	2	US-08-488-379-38	Sequence 38, Appl
15	30	76.9	20	4	US-08-475-399A-38	Sequence 38, Appl
16	30	76.9	20	5	PCT-US93-07545-38	Sequence 38, Appl
17	30	76.9	141	4	US-09-107-532A-6346	Sequence 6346, Ap
18	30	76.9	180	4	US-09-540-236-3016	Sequence 3016, Ap
19	30	76.9	244	3	US-09-003-287-6	Sequence 6, Appli
20	30	76.9	244	3	US-09-003-287-8	Sequence 8, Appli
21	30	76.9	244	3	US-09-518-988-2	Sequence 2, Appli
22	30	76.9	362	2	US-09-080-897-6	Sequence 6, Appli
23	30	76.9	362	3	US-09-323-735-6	Sequence 6, Appli
24	30	76.9	459	4	US-09-543-681A-6287	Sequence 6287, Ap
25	30	76.9	1250	3	US-08-938-291A-9	Sequence 9, Appli
26	30	76.9	1250	4	US-09-589-619-9	Sequence 9, Appli
27	29	74.4	9	1	US-08-146-145-6	Sequence 6, Appli
28	29	74.4	127	3	US-08-467-023-189	Sequence 189, App
29	29	74.4	204	4	US-09-489-039A-8937	Sequence 8937, Ap
30	29	74.4	475	4	US-09-252-991A-32806	Sequence 32806, A
31	29	74.4	514	3	US-08-467-023-134	Sequence 134, App
32	29	74.4	531	4	US-09-489-039A-12406	Sequence 12406, A
33	29	74.4	574	3	US-09-385-028-4	Sequence 4, Appli
34	29	74.4	574	4	US-09-726-614-4	Sequence 4, Appli
35	29	74.4	574	4	US-09-385-040-4	Sequence 4, Appli
36	29	74.4	602	2	US-08-419-652-6	Sequence 6, Appli
37	29	74.4	607	4	US-09-134-001C-2994	Sequence 2994, Ap
38	29	74.4	617	4	US-09-328-352-6700	Sequence 6700, Ap
39	29	74.4	698	3	US-08-941-445A-11	Sequence 11, Appl
40	29	74.4	771	1	US-07-923-976-6	Sequence 6, Appli
41	29	74.4	783	6	5422248-2	Patent No. 5422248
42	29	74.4	836	1	US-07-923-976-4	Sequence 4, Appli
43	29	74.4	863	1	US-07-923-976-8	Sequence 8, Appli
44	29	74.4	900	4	US-09-252-991A-25011	Sequence 25011, A
45	28	71.8	76	4	US-09-621-976-6895	Sequence 6895, Ap

ALIGNMENTS

RESULT 1
US-09-543-681A-4935
; Sequence 4935, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4935
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4935

Query Match 84.6%; Score 33; DB 4; Length 485;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMX 9
              :|||::||:
Db        457 TLYESIGMA 465
```

RESULT 2
US-08-272-255-8
; Sequence 8, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/272,255
;      FILING DATE:  08-JUL-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leary Ph.D., Kathryn
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  UPN-1795
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  566 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-272-255-8

  Query Match             84.6%;  Score 33;  DB 2;  Length 566;
  Best Local Similarity   55.6%;  Pred. No. 78;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||:|||::
Db         87 RLYDNVGLY 95


RESULT 3
PCT-US95-08565-8
; Sequence 8, Application PC/TUS9508565
;  GENERAL INFORMATION:
;    APPLICANT:  Cashmore, Anthony R.
;    APPLICANT:  Ahmad, Margaret
;    APPLICANT:  Lin, Chentao
;    TITLE OF INVENTION:  Blue Light Photoreceptors and Methods of
;    TITLE OF INVENTION:  Using the Same
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz & Norris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/08565
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/272,255
;      FILING DATE:  08-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leary Ph.D., Kathryn
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  UPN-1795
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  566 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-08565-8

  Query Match             84.6%;  Score 33;  DB 5;  Length 566;
  Best Local Similarity   55.6%;  Pred. No. 78;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||:|||::
Db         87 RLYDNVGLY 95


RESULT 4
US-08-202-389-12
; Sequence 12, Application US/08202389
; Patent No. 5536636
;  GENERAL INFORMATION:
;    APPLICANT:  Freeman Jr., Robert M.
;    APPLICANT:  Plutzky, Jorge
;    APPLICANT:  Neel, Benjamin G.
;    APPLICANT:  Rosenberg, Robert D.
;    TITLE OF INVENTION:  IDENTIFICATION OF NOVEL TYROSINE
;    TITLE OF INVENTION:  PHOSPHATASES HAVING SH2 DOMAINS
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/202,389
;      FILING DATE:  28-FEB-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/983,926
;      FILING DATE:  01-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/829,141
;      FILING DATE:  31-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/721,112
;      FILING DATE:  26-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  BIH92-05MA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  593 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-202-389-12

  Query Match             84.6%;  Score 33;  DB 1;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 82;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        578 RVYENVGLM 586
```

```
RESULT 5
US-08-018-129-5
; Sequence 5, Application US/08018129
; Patent No. 5589375
;  GENERAL INFORMATION:
;    APPLICANT:  Ullrich, Axel
;    APPLICANT:  Vogel, Wolfgang
;    TITLE OF INVENTION:  PTP 1D: A NOVEL PROTEIN TYROSINE
;    TITLE OF INVENTION:  PHOSPHATASE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/018,129
;      FILING DATE:  19930216
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7683-017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  593 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-018-129-5

  Query Match             84.6%;  Score 33;  DB 1;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 82;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        578 RVYENVGLM 586


RESULT 6
US-08-448-250-5
; Sequence 5, Application US/08448250
; Patent No. 5981251
;  GENERAL INFORMATION:
;    APPLICANT:  Ullrich, Axel
;    APPLICANT:  Vogel, Wolfgang
;    TITLE OF INVENTION:  PTP 1D: A NOVEL PROTEIN TYROSINE
;    TITLE OF INVENTION:  PHOSPHATASE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,250
;      FILING DATE:  23-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/018,129
;      FILING DATE:  16-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7683-017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  593 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-448-250-5

  Query Match             84.6%;  Score 33;  DB 2;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 82;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        578 RVYENVGLM 586


RESULT 7
US-09-282-257-5
; Sequence 5, Application US/09282257
; Patent No. 6548641
;  GENERAL INFORMATION:
;    APPLICANT:  Ullrich, Axel
;    APPLICANT:  Vogel, Wolfgang
;    TITLE OF INVENTION:  PTP 1D: A NOVEL PROTEIN TYROSINE
;    TITLE OF INVENTION:  PHOSPHATASE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/282,257
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/018,129
;      FILING DATE:  16-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7683-017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  593 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-282-257-5

  Query Match             84.6%;  Score 33;  DB 4;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 82;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        578 RVYENVGLM 586


RESULT 8
US-09-489-039A-12637
; Sequence 12637, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 12637
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12637

  Query Match             82.1%;  Score 32;  DB 4;  Length 214;
  Best Local Similarity   66.7%;  Pred. No. 42;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|| ||||:
Db        148 SLYGNVGMA 156


RESULT 9
US-09-134-001C-3551
; Sequence 3551, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3551
;   LENGTH: 325
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3551

  Query Match             79.5%;  Score 31;  DB 4;  Length 325;
  Best Local Similarity   44.4%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::||::||:
Db        167 QIYESIGMD 175


RESULT 10
US-09-134-000C-3814
; Sequence 3814, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3814
;   LENGTH: 335
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-3814

  Query Match             79.5%;  Score 31;  DB 4;  Length 335;
  Best Local Similarity   55.6%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||:| ||:
Db         53 LLYKNTGMT 61


RESULT 11
US-08-788-674-4
; Sequence 4, Application US/08788674
; Patent No. 5922315
;  GENERAL INFORMATION:
;    APPLICANT:  Roy, Soumitra
;    TITLE OF INVENTION:  Adenoviruses Having Altered
;    TITLE OF INVENTION:  Hexon Proteins
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain,
;      ADDRESSEE:  Gilfillan, Cecchi, Stewart &
;      ADDRESSEE:  Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/788,674
;      FILING DATE:  24-JAN-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Olstein, Elliot M.
;      REGISTRATION NUMBER:  24,025
;      REFERENCE/DOCKET NUMBER:  271010-363
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  973-994-1700
;      TELEFAX:  973-994-1744
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  919 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  predicted hexon protein sequence
;      NAME/KEY:  for human Adenovirus 12
US-08-788-674-4

  Query Match             79.5%;  Score 31;  DB 2;  Length 919;
  Best Local Similarity   55.6%;  Pred. No. 3.5e+02;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|| |||::
Db        439 FLYSNVGLY 447


RESULT 12
US-09-376-343-3
; Sequence 3, Application US/09376343
; Patent No. 6506592
; GENERAL INFORMATION:
;  APPLICANT: Blum, Paul H.
;  TITLE OF INVENTION: Hyperthermophilic Alpha-Glucosidase Gene and Its Use
;  FILE REFERENCE: N1231-200
;  CURRENT APPLICATION NUMBER: US/09/376,343
;  CURRENT FILING DATE: 1999-08-18
;  EARLIER APPLICATION NUMBER: 60/096,860
;  EARLIER FILING DATE: 1998-08-18
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Sulfolobus solfataricus
US-09-376-343-3

  Query Match             76.9%;  Score 30;  DB 4;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 7.3;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||:|::
Db          5 KIYENLGVY 13


RESULT 13
US-08-480-190-38
; Sequence 38, Application US/08480190
; Patent No. 5827516
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,190
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-480-190-38

  Query Match             76.9%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   71.4%;  Pred. No. 7.8;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||:|||
Db          2 TLYQNVG 8


RESULT 14
US-08-488-379-38
; Sequence 38, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,379
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  38:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-488-379-38

  Query Match             76.9%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   71.4%;  Pred. No. 7.8;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||:|||
Db          2 TLYQNVG 8


RESULT 15
US-08-475-399A-38
; Sequence 38, Application US/08475399A
; Patent No. 6509033
;  GENERAL INFORMATION:
;    APPLICANT:  Urban, Robert G.
;    APPLICANT:  Chicz, Roman M.
;    APPLICANT:  Vignali, Dario A.A.
;    APPLICANT:  Hedley, Mary L.
;    APPLICANT:  Stern, Lawrence J.
;    APPLICANT:  Strominger, Jack L.
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  276
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,399A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  15-JUN-1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  11-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  00246/168003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-507
;      TELEFAX:  617/542-890
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-475-399A-38

  Query Match             76.9%;  Score 30;  DB 4;  Length 20;
  Best Local Similarity   71.4%;  Pred. No. 7.8;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||:|||
Db          2 TLYQNVG 8


Search completed: July 15, 2004, 07:31:19
Job time : 15.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:27:08 ; Search time 40 Seconds
(without alignments)
70.326 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOM8.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOM8.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	100.0	9	10	US-09-998-350-1	Sequence 1, Appli
2	39	100.0	9	10	US-09-998-350-3	Sequence 3, Appli
3	39	100.0	9	10	US-09-998-350-7	Sequence 7, Appli
4	39	100.0	10	10	US-09-998-350-4	Sequence 4, Appli
5	39	100.0	10	10	US-09-998-350-5	Sequence 5, Appli
6	39	100.0	10	10	US-09-998-350-6	Sequence 6, Appli
7	39	100.0	10	10	US-09-998-350-8	Sequence 8, Appli
8	39	100.0	10	10	US-09-998-350-11	Sequence 11, Appl
9	39	100.0	10	10	US-09-998-350-14	Sequence 14, Appl
10	39	100.0	11	14	US-10-013-815-32	Sequence 32, Appl
11	39	100.0	26	10	US-09-998-350-18	Sequence 18, Appl
12	39	100.0	26	10	US-09-998-350-19	Sequence 19, Appl
13	34	87.2	134	16	US-10-437-963-168439	Sequence 168439,
14	33	84.6	10	10	US-09-998-350-10	Sequence 10, Appl
15	33	84.6	10	10	US-09-998-350-12	Sequence 12, Appl
16	33	84.6	10	10	US-09-998-350-13	Sequence 13, Appl
17	33	84.6	144	12	US-10-424-599-278276	Sequence 278276,
18	33	84.6	474	15	US-10-369-493-21087	Sequence 21087, A
19	33	84.6	475	12	US-10-282-122A-68739	Sequence 68739, A
20	33	84.6	475	15	US-10-369-493-411	Sequence 411, App
21	33	84.6	589	16	US-10-322-281-270	Sequence 270, App
22	33	84.6	593	9	US-09-920-021A-3	Sequence 3, Appli
23	33	84.6	593	14	US-10-262-552-2	Sequence 2, Appli
24	33	84.6	593	15	US-10-366-547-16	Sequence 16, Appl
25	33	84.6	593	15	US-10-366-547-26	Sequence 26, Appl
26	33	84.6	593	15	US-10-366-547-28	Sequence 28, Appl
27	33	84.6	593	15	US-10-366-547-30	Sequence 30, Appl
28	33	84.6	593	15	US-10-366-547-32	Sequence 32, Appl
29	33	84.6	593	16	US-10-444-795B-789	Sequence 789, App
30	33	84.6	593	16	US-10-444-795B-791	Sequence 791, App
31	33	84.6	593	16	US-10-703-210-2	Sequence 2, Appli
32	33	84.6	597	14	US-10-038-010-22	Sequence 22, Appl
33	33	84.6	597	15	US-10-366-547-14	Sequence 14, Appl
34	33	84.6	612	16	US-10-322-281-267	Sequence 267, App
35	33	84.6	1096	14	US-10-128-714-3376	Sequence 3376, Ap
36	33	84.6	1096	14	US-10-128-714-8376	Sequence 8376, Ap
37	32	82.1	606	15	US-10-369-493-13804	Sequence 13804, A
38	32	82.1	1900	9	US-09-801-368-390	Sequence 390, App
39	31	79.5	68	12	US-10-424-599-259481	Sequence 259481,
40	31	79.5	139	12	US-10-377-097-55	Sequence 55, Appl
41	31	79.5	149	12	US-10-282-122A-62320	Sequence 62320, A
42	31	79.5	149	12	US-10-282-122A-64762	Sequence 64762, A
43	31	79.5	149	14	US-10-080-170-563	Sequence 563, App
44	31	79.5	149	16	US-10-080-170-563	Sequence 563, App
45	31	79.5	187	12	US-10-424-599-172341	Sequence 172341,

ALIGNMENTS

RESULT 1
US-09-998-350-1
; Sequence 1, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is an amide, i.e. C(O)NH
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (1)..(9)
;   OTHER INFORMATION: Xaa (Gla) and Tyr at position 9 are bridged together, making this
;   OTHER INFORMATION: peptide cyclic
US-09-998-350-1

  Query Match             100.0%;  Score 39;  DB 10;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.2e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 2
US-09-998-350-3
; Sequence 3, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa is any amino acid other than Glu
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(9)
;   OTHER INFORMATION: Xaa and Tyr at position 9 are bridged together, making this pepti
;   OTHER INFORMATION: de cyclic
US-09-998-350-3

  Query Match             100.0%;  Score 39;  DB 10;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.2e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 3
US-09-998-350-7
; Sequence 7, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa has a ClCH2C(O)- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 has a -C(CH2SH)C(O)NH2 group attached
US-09-998-350-7

  Query Match             100.0%;  Score 39;  DB 10;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.2e+06;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 4
US-09-998-350-4
; Sequence 4, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys at position 10 is an amide, i.e., C(O)NH
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Xaa (Gla) and Cys are bridged together, making this peptide cycli
;   OTHER INFORMATION: c
US-09-998-350-4

  Query Match             100.0%;  Score 39;  DB 10;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.24;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 5
US-09-998-350-5
; Sequence 5, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
US-09-998-350-5

  Query Match             100.0%;  Score 39;  DB 10;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.24;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 6
US-09-998-350-6
; Sequence 6, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla(OtBu)2, which is di- tert-butoxy-gamma-carboxy-L-glutam
;   OTHER INFORMATION: ic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: Tyr at position 3 is modified to Tyr(tBu), which is tert-butyl-ty
;   OTHER INFORMATION: rosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is is trytyl-asp
;   OTHER INFORMATION: aragine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is modified to Tyr(tBu), which is tert-butyl-ty
;   OTHER INFORMATION: rosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys at position 10 is modified to Cys(Trt), which is trytyl-cyste
;   OTHER INFORMATION: ine, and Cys(Trt) is connected to a resin
US-09-998-350-6

  Query Match             100.0%;  Score 39;  DB 10;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.24;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 7
US-09-998-350-8
; Sequence 8, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
```

```
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Adi, which is alpha-amino-adipic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa has a CH2CO- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Xaa (Adi) and Cys are bridged together, making this peptide cycli
;   OTHER INFORMATION: c
US-09-998-350-8

  Query Match             100.0%;  Score 39;  DB 10;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.24;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 8
US-09-998-350-11
; Sequence 11, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: Tyr at position 3 is modified to Tyr(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: tyrosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trityl-aspara
;   OTHER INFORMATION: gine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: tyrosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa = Nle, which is norleucine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa is an amide and is attached to a resin
US-09-998-350-11

  Query Match             100.0%;  Score 39;  DB 10;  Length 10;
  Best Local Similarity   77.8%;  Pred. No. 0.24;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          1 ELYENVGMY 9


RESULT 9
US-09-998-350-14
; Sequence 14, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: glutamic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trytyl-aspara
;   OTHER INFORMATION: gine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
;   OTHER INFORMATION: tyrosine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa = Adi(OAl), which is allyloxy-alpha-amino-adipic acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Xaa is an amide, i.e., C(O)NH
US-09-998-350-14

  Query Match             100.0%;  Score 39;  DB 10;  Length 10;
  Best Local Similarity   77.8%;  Pred. No. 0.24;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          1 ELYENVGMY 9


RESULT 10
US-10-013-815-32
; Sequence 32, Application US/10013815
; Publication No. US20030105000A1
; GENERAL INFORMATION:
;  APPLICANT: Pero, Stephanie
;  APPLICANT:  Krag, David
;  APPLICANT:  Oligino, Lyn
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING GRB7
;  FILE REFERENCE: V0139/7048 (HCL/MAT)
;  CURRENT APPLICATION NUMBER: US/10/013,815
;  CURRENT FILING DATE:  2001-11-05
;  PRIOR APPLICATION NUMBER: US 60/245,755
;  PRIOR FILING DATE: 2000-11-03
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20030105000A1-phosphorylated peptide with YEN motif
US-10-013-815-32

  Query Match             100.0%;  Score 39;  DB 14;  Length 11;
  Best Local Similarity   77.8%;  Pred. No. 0.27;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||||:
Db          2 ELYENVGMY 10


RESULT 11
US-09-998-350-18
; Sequence 18, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
US-09-998-350-18

  Query Match             100.0%;  Score 39;  DB 10;  Length 26;
  Best Local Similarity   88.9%;  Pred. No. 0.68;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 12
US-09-998-350-19
; Sequence 19, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND M
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa (G1a) has a CH2CO- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
US-09-998-350-19

  Query Match             100.0%;  Score 39;  DB 10;  Length 26;
  Best Local Similarity   88.9%;  Pred. No. 0.68;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ||||||||:
Db          1 XLYENVGMY 9


RESULT 13
US-10-437-963-168439
; Sequence 168439, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168439
;   LENGTH: 134
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(134)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_66953C.1.pep
US-10-437-963-168439

  Query Match             87.2%;  Score 34;  DB 16;  Length 134;
  Best Local Similarity   55.6%;  Pred. No. 42;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||:||:
Db        103 DIYENMGMK 111


RESULT 14
US-09-998-350-10
; Sequence 10, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Glu has a CH2CO- group attached
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: Xaa = Nle, which is norleucine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Glu and Cys are bridged together, making this peptide cyclic
US-09-998-350-10

  Query Match             84.6%;  Score 33;  DB 10;  Length 10;
  Best Local Similarity   85.7%;  Pred. No. 4.1;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||||
Db          1 ELYENVG 7


RESULT 15
US-09-998-350-12
; Sequence 12, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
;  APPLICANT: Roller, Peter P
;  APPLICANT:  Long, Ya-Qiu
;  APPLICANT:  Lung, Feng-Di T
;  APPLICANT:  King, Richter C
;  APPLICANT:  Yang, Dajun
;  TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
;  TITLE OF INVENTION:  BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
;  TITLE OF INVENTION:  SYNTHESIS AND USE
;  FILE REFERENCE: 214683
;  CURRENT APPLICATION NUMBER: US/09/998,350
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: PCT/US00/15201
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: 60/137,187
;  PRIOR FILING DATE: 1999-06-02
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: Xaa = Nle, which is norleucine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Glu at position 1 and Cys are bridged together, making this pepti
;   OTHER INFORMATION: de cyclic
US-09-998-350-12

  Query Match             84.6%;  Score 33;  DB 10;  Length 10;
  Best Local Similarity   85.7%;  Pred. No. 4.1;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||||
Db          1 ELYENVG 7


Search completed: July 15, 2004, 07:32:50
Job time : 41 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:23:22 ; Search time 11.5 Seconds
(without alignments)
75.280 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	87.2	688	2	T33708	hypothetical prote
2	33	84.6	565	2	S67298	deoxyribodipyrimid
3	33	84.6	593	1	JN0805	protein-tyrosine-p
4	33	84.6	593	2	JC5167	protein-tyrosine-p
5	33	84.6	595	1	A55651	protein-tyrosine-p
6	33	84.6	597	1	A53593	protein-tyrosine-p
7	33	84.6	700	2	T20550	hypothetical prote
8	32	82.1	178	2	B69944	hypothetical prote
9	32	82.1	231	2	H85138	hypothetical prote
10	32	82.1	352	2	D72264	hypothetical prote
11	32	82.1	367	2	AD1786	cell division prot
12	32	82.1	369	2	AF1410	cell division prot
13	32	82.1	617	2	C97152	conjugative transf
14	32	82.1	1364	1	AJFFPP	phosphoribosylamin
15	32	82.1	1900	2	S45530	probable 1-phospha
16	31	79.5	99	2	S44632	f22b7.3 protein -
17	31	79.5	149	2	C70878	hypothetical prote
18	31	79.5	201	2	T25706	hypothetical prote
19	31	79.5	219	2	E75143	phosphoglycolate p
20	31	79.5	307	2	T34973	5,10-methylenetetr
21	31	79.5	356	2	C97265	mannose-1-phosphat
22	31	79.5	360	2	S57777	cysteine proteinas
23	31	79.5	402	2	T13614	N-acetyltransferas
24	31	79.5	406	2	D86895	membrane protein [
25	31	79.5	432	2	F83903	hypothetical prote
26	31	79.5	454	2	C86766	hypothetical prote
27	31	79.5	468	2	S37217	hexon protein - hu
28	31	79.5	526	1	VGVNSG	spike glycoprotein
29	31	79.5	533	2	H71492	probable hsp-60 -
30	31	79.5	739	2	T45429	polyphosphate kina
31	31	79.5	742	2	E70673	probable ppk prote
32	31	79.5	919	2	S33942	hexon protein - hu
33	31	79.5	1302	2	A41249	multidrug resistan
34	30	76.9	20	2	PL0161	hemagglutinin - In
35	30	76.9	159	2	E86760	conserved hypothet
36	30	76.9	160	2	AB3559	transcription regu
37	30	76.9	164	2	B87032	conserved hypothet
38	30	76.9	174	2	D81822	MutT-like protein
39	30	76.9	174	2	E81055	MutT/nudix family
40	30	76.9	178	2	S72948	hypothetical prote
41	30	76.9	193	2	G82294	MutT/nudix family
42	30	76.9	205	2	D44583	venom allergen ant
43	30	76.9	240	2	H98031	hypothetical prote
44	30	76.9	244	2	B95166	ABC transporter, A
45	30	76.9	244	2	A39365	cyanamide hydratas

ALIGNMENTS

RESULT 1
T33708
hypothetical protein F58E2.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000
C;Accession: T33708
R;Goela, D.; Delehaunty, A.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of C. elegans cosmid F58E2.
A;Reference number: Z21390
A;Accession: T33708
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-688 <GOE>
A;Cross-references: EMBL:AF100659; PIDN:AAC68967.1; GSPDB:GN00022; CESP:F58E2.4
A;Experimental source: strain Bristol N2; clone F58E2
C;Genetics:
A;Gene: CESP:F58E2.4
A;Map position: 4
A;Introns: 228/3; 309/3; 344/2; 602/3
C;Superfamily: Caenorhabditis elegans hypothetical protein F58E2.3

Query Match 87.2%; Score 34; DB 2; Length 688;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XLYENVGMX 9
              ::||||::
Db        393 LIYENVGLS 401
```

RESULT 2
S67298
deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O6771; protein YOR386w
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C;Accession: S67298; A23964; A24046
R;Delius, H.; Hebling, U.; Hofmann, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67261
A;Accession: S67298
A;Molecule type: DNA
A;Residues: 1-565 <DEL>
A;Cross-references: EMBL:Z75294; NID:g1420830; PIDN:CAA99718.1; PID:g1420831; MIPS:YOR386
A;Experimental source: strain S288C
R;Yasui, A.; Langeveld, S.A.
Gene 36, 349-355, 1985
A;Title: Homology between the photoreactivation genes of Saccharomyces cerevisiae and Esc
A;Reference number: A23964; MUID:86083177; PMID:3000886
A;Accession: A23964
A;Molecule type: DNA

A;Residues: 1-76,'A',78-164,'S',166-168,'T',170-199,'S',201-350,'R',352-364,'E',366-472,
A;Cross-references: EMBL:M11578; NID:g172169; PIDN:AAA34875.1; PID:g172170
R;Sancar, G.B.
Nucleic Acids Res. 13, 8231-8246, 1985
A;Title: Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of the PHR1 pho
A;Reference number: A24046; MUID:86067229; PMID:3906569
A;Accession: A24046
A;Molecule type: DNA
A;Residues: 1-565 <SAN>
A;Cross-references: EMBL:X03183; NID:g4175; PIDN:CAA26944.1; PID:g4176
C;Genetics:
A;Gene: SGD:PHR1
A;Cross-references: SGD:S0005913; MIPS:YOR386w
A;Map position: 15R
C;Superfamily: deoxyribodipyrimidine photo-lyase
C;Keywords: carbon-carbon lyase

Query Match 84.6%; Score 33; DB 2; Length 565;
Best Local Similarity 55.6%; Pred. No. 44;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
:||:|||::
Db 86 RLYDNVGLY 94

RESULT 3
JN0805
protein-tyrosine-phosphatase (EC 3.1.3.48) PTPN11, nonreceptor type 11 [validated] - hum
N;Alternate names: BPTP-3; protein-tyrosine-phosphatase SHP-2; PTP1D; PTP2C; SH-PTP2; SH
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1994 #sequence_revision 19-May-1994 #text_change 21-Jun-2002
C;Accession: JN0805; A46210; A47386; A47244; S27398; C44929; S31767
R;Bastien, L.; Ramachandran, C.; Liu, S.; Adam, M.
Biochem. Biophys. Res. Commun. 196, 124-133, 1993
A;Title: Cloning, expression and mutational analysis of SH-PTP2, human protein-tyrosine
A;Reference number: JN0805; MUID:94029983; PMID:8216283
A;Accession: JN0805
A;Molecule type: mRNA
A;Residues: 1-593 <BAS>
A;Cross-references: GB:L07527; NID:g292406; PIDN:AAA17022.1; PID:g292407
R;Vogel, W.; Lammers, R.; Huang, J.; Ullrich, A.
Science 259, 1611-1614, 1993
A;Title: Activation of a phosphotyrosine phosphatase by tyrosine phosphorylation.
A;Reference number: A46210; MUID:93206095; PMID:7681217
A;Accession: A46210
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-593 <VOG>
A;Cross-references: EMBL:X70766; NID:g35783; PIDN:CAA50045.1; PID:g35784
A;Experimental source: SK-BR-3 mammary carcinoma cells
A;Note: sequence extracted from NCBI backbone (NCBIP:127775)
R;Ahmad, S.; Banville, D.; Zhao, Z.; Fischer, E.H.; Shen, S.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 2197-2201, 1993
A;Title: A widely expressed human protein-tyrosine phosphatase containing src homology 2
A;Reference number: A47386; MUID:93211929; PMID:7681589
A;Accession: A47386
A;Molecule type: mRNA
A;Residues: 1-593 <AHM>
A;Experimental source: umbilical cord
A;Note: sequence extracted from NCBI backbone (NCBIN:128129, NCBIP:128131)
R;Freeman Jr., R.M.; Plutzky, J.; Neel, B.G.
Proc. Natl. Acad. Sci. U.S.A. 89, 11239-11243, 1992
A;Title: Identification of a human src homology 2-containing protein-tyrosine-phosphatas
A;Reference number: A47244; MUID:93087502; PMID:1280823
A;Accession: A47244
A;Molecule type: mRNA
A;Residues: 1-593 <FRE>
A;Cross-references: GB:L03535; NID:g338081; PIDN:AAA36611.1; PID:g338082
A;Note: sequence extracted from NCBI backbone (NCBIN:119760, NCBIP:119761)
R;Adachi, M.; Sekiya, M.; Miyachi, T.; Matsuno, K.; Hinoda, Y.; Imai, K.; Yachi, A.
FEBS Lett. 314, 335-339, 1992
A;Title: Molecular cloning of a novel protein-tyrosine phosphatase SH-PTP3 with sequence
A;Reference number: S27398; MUID:93106179; PMID:1281790
A;Accession: S27398
A;Molecule type: mRNA
A;Residues: 1-534,'R',536-547,'P',549-593 <AD2>
A;Cross-references: DDBJ:D13540; NID:g220071; PIDN:BAA02740.2; PID:g4519425
R;Adachi, M.; Sekiya, M.; Arimura, Y.; Takekawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yach
Cancer Res. 52, 737-740, 1992
A;Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6.
A;Reference number: A44929; MUID:92119637; PMID:1370651
A;Accession: C44929
A;Molecule type: mRNA
A;Residues: 'Q',370-460 <ADA>
A;Cross-references: GB:S78088; NID:g243547; PIDN:AAB21148.1; PID:g243548
A;Experimental source: pre-B cell NALM-6
A;Note: sequence extracted from NCBI backbone (NCBIN:78088, NCBIP:78089)
A;Note: the authors did not report the entire codon for residue 92
C;Comment: This ubiquitous enzyme plays a critical role in regulating physiological cell
C;Genetics:
A;Gene: GDB:PTPN11
A;Cross-references: GDB:137093; OMIM:176876
A;Map position: 12q24.1-12q24.1
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phospha
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F;6-100/Domain: SH2 homology <SH2A>
F;112-214/Domain: SH2 homology <SH2B>
F;273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F;459/Active site: Cys (phosphocysteine intermediate) #status predicted
F;465/Binding site: substrate phosphate (Arg) #status predicted

Query Match 84.6%; Score 33; DB 1; Length 593;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
::|||||::
Db 578 RVYENVGLM 586

RESULT 4
JC5167
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - chicken
N;Alternate names: phosphotyrosine phosphatase; PTP1D; PTP2c; SH-PTP2; SyP
C;Species: Gallus gallus (chicken)
C;Date: 21-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 16-Jul-1999
C;Accession: JC5167
R;Park, C.Y.; LaMontagne, K.R.; Tonks, N.K.; Hayman, M.J.
Gene 177, 93-97, 1996
A;Title: Cloning and expression of the chicken protein tyrosine phosphatase SH-PTP2.
A;Reference number: JC5167; MUID:97080506; PMID:8921851
A;Contents: erythroblast
A;Accession: JC5167
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-593 <PAR>
A;Cross-references: GB:U38620; NID:g1054939; PIDN:AAC60049.1; PID:g1054940
C;Comment: This enzyme plays positive roles in mitogenic signaling and early development.
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phospha
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F;6-105,112-193/Domain: SH2 #status predicted <SH2>
F;6-100/Domain: SH2 homology <SH2A>
F;112-214/Domain: SH2 homology <SH2B>
F;273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F;559-570/Region: proline-rich
F;459/Active site: Cys (phosphocysteine intermediate) #status predicted
F;465/Binding site: substrate phosphate (Arg) #status predicted
F;542,547,580/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 84.6%; Score 33; DB 2; Length 593;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XLYENVGMX 9
::|||||::

```
Db        578 RVYENVGLM 586


RESULT 5
A55651
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - African clawed frog
N;Alternate names: SH-PTP2
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A55651
R;Tang, T.L.; Freeman Jr., R.M.; O'Reilly, A.M.; Neel, B.G.; Sokol, S.Y.
Cell 80, 473-483, 1995
A;Title: The SH2-containing protein-tyrosine phosphatase SH-PTP2 is required upstream of
A;Reference number: A55651; MUID:95163101; PMID:7859288
A;Accession: A55651
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-595 <TAN>
A;Cross-references: GB:U15287; NID:g601781; PIDN:AAA65731.1; PID:g601782
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosph
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatas
F;6-100/Domain: SH2 homology <SH2A>
F;112-214/Domain: SH2 homology <SH2B>
F;273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
F;459/Active site: Cys (phosphocysteine intermediate) #status predicted
F;465/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             84.6%;  Score 33;  DB 1;  Length 595;
  Best Local Similarity   55.6%;  Pred. No. 46;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        580 RVYENVGLL 588


RESULT 6
A53593
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - rat
N;Alternate names: PTPase L1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A53593; S29281
R;Mei, L.; Doherty, C.A.; Huganir, R.L.
J. Biol. Chem. 269, 12254-12262, 1994
A;Title: RNA splicing regulates the activity of a SH2 domain-containing protein tyrosine
A;Reference number: A53593; MUID:94216346; PMID:7512964
A;Accession: A53593
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-597 <MEI>
A;Cross-references: GB:U05963; NID:g458332; PIDN:AAA19133.1; PID:g458333
R;Hiraga, A.; Munakata, H.; Hata, K.; Suzuki, Y.; Tsuiki, S.
Eur. J. Biochem. 209, 195-206, 1992
A;Title: Purification and characterization of a rat liver protein-tyrosine phosphatase w
A;Reference number: S29281; MUID:93011127; PMID:1382983
A;Accession: S29281
A;Molecule type: protein
A;Residues: 24-31;36-54;56-89;100-103,'X',105-108;'X',113-120;132-155;179-198;214-233;24
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-phosph
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosi
F;6-100/Domain: SH2 homology <SH2A>
F;112-214/Domain: SH2 homology <SH2B>
F;273-514/Domain: protein-tyrosine-phosphatase homology <PTP>
F;463/Active site: Cys (phosphocysteine intermediate) #status predicted
F;469/Binding site: substrate phosphate (Arg) #status predicted

  Query Match             84.6%;  Score 33;  DB 1;  Length 597;
  Best Local Similarity   55.6%;  Pred. No. 46;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        582 RVYENVGLM 590


RESULT 7
T20550
hypothetical protein F07C6.4b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T20550; T23678
R;Steward, C.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19290
A;Accession: T20550
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-700 <WIL>
A;Cross-references: EMBL:Z69659; PIDN:CAA93486.1; GSPDB:GN00022; CESP:F07C6.4b
A;Experimental source: clone F07C6
R;Lightning, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19780
A;Accession: T23678
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-700 <WI2>
A;Cross-references: EMBL:Z81102; PIDN:CAB03204.1; GSPDB:GN00022; CESP:F07C6.4b
A;Experimental source: clone M02B1
C;Genetics:
A;Gene: CESP:F07C6.4b
A;Map position: 4
A;Introns: 21/3; 58/2; 111/1; 159/3; 195/3; 272/1; 328/2; 399/2; 423/3; 546/3; 564/1; 61

  Query Match             84.6%;  Score 33;  DB 2;  Length 700;
  Best Local Similarity   85.7%;  Pred. No. 56;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||||
Db        398 PLYENVG 404


RESULT 8
B69944
hypothetical protein yqaC - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: B69944
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: B69944
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-178 <KUN>
A;Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14578.1; PID:e1183866;
A;Experimental source: strain 168
C;Genetics:
A;Gene: yqaC

  Query Match             82.1%;  Score 32;  DB 2;  Length 178;
```

```
  Best Local Similarity   55.6%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||:| ||:
Db        129 SLYDNAGME 137


RESULT 9
H85138
hypothetical protein AT4g12900 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C;Accession: H85138
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: H85138
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-231 <STO>
A;Cross-references: GB:NC_001268; NID:g7267992; PIDN:CAB78332.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g12900
A;Map position: 4
C;Superfamily: Arabidopsis thaliana hypothetical protein F7A7.100

  Query Match             82.1%;  Score 32;  DB 2;  Length 231;
  Best Local Similarity   71.4%;  Pred. No. 26;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||:|
Db        182 PLYENIG 188


RESULT 10
D72264
hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 28-Jul-2000
C;Accession: D72264
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72264
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <ARN>
A;Cross-references: GB:AE001789; GB:AE000512; NID:g4981904; PIDN:AAD36419.1; PID:g498190
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1348
C;Superfamily: Thermotoga maritima hypothetical protein TM1348

  Query Match             82.1%;  Score 32;  DB 2;  Length 352;
  Best Local Similarity   55.6%;  Pred. No. 42;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||| :||:
Db        330 RLYEEIGMH 338


RESULT 11
AD1786
cell division protein FtsW homolog lin2834 [imported] - Listeria innocua (strain Clip112
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AD1786
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1786
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-367 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC98060.1; PID:g16415369; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2834
C;Superfamily: rod shape-determining protein

  Query Match             82.1%;  Score 32;  DB 2;  Length 367;
  Best Local Similarity   44.4%;  Pred. No. 44;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :::||:||:
Db        316 NIFENIGMT 324


RESULT 12
AF1410
cell division protein FtsW homolog lmo2687 [imported] - Listeria monocytogenes (strain EC
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AF1410
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1410
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-369 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00900.1; PID:g16412187; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2687
C;Superfamily: rod shape-determining protein

  Query Match             82.1%;  Score 32;  DB 2;  Length 369;
  Best Local Similarity   44.4%;  Pred. No. 44;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :::||:||:
Db        316 NIFENIGMT 324


RESULT 13
C97152
conjugative transfer gene TrsE homolog, ATPase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: C97152
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clos
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97152
```

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-617 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80006.1; PID:g15025033; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2047

  Query Match             82.1%;  Score 32;  DB 2;  Length 617;
  Best Local Similarity   55.6%;  Pred. No. 80;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||||:|::
Db        369 QLYENLGIT 377


RESULT 14
AJFFPP
phosphoribosylamine-glycine ligase (EC 6.3.4.13) - fruit fly (Drosophila pseudoobscura)
N;Alternate names: glycinamide ribonucleotide synthetase (GARSase); glycinamide ribonucl
N;Contains: phosphoribosylamine-glycine ligase (EC 6.3.4.13); phosphoribosylformylglycin
C;Species: Drosophila pseudoobscura
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Jun-2002
C;Accession: S01204
R;Henikoff, S.; Eghtedarzadeh, M.K.
Genetics 117, 711-725, 1987
A;Title: Conserved arrangement of nested genes at the Drosophila Gart locus.
A;Reference number: S01204; MUID:88112752; PMID:3123310
A;Accession: S01204
A;Molecule type: DNA
A;Residues: 1-1364 <HEN>
A;Cross-references: EMBL:X06285; NID:g9055; PIDN:CAA29611.1; PID:g295787
A;Note: monofunctional phosphoribosylamine-glycine ligase, prepared by alternative splic
C;Genetics:
A;Gene: Gart
A;Cross-references: FlyBase:FBgn0000053
A;Map position: 4 88
A;Introns: 59/3; 142/2; 359/1; 434/2; 575/1; 927/1
C;Superfamily: Drosophila purine synthesis multifunctional protein; phosphoribosylamine-
myltransferase homology
C;Keywords: alternative splicing; cyclo-ligase; methyltransferase; multifunctional enzym
F;4-430/Domain: phosphoribosylamine-glycine ligase homology <PGL>
F;444-775/Domain: phosphoribosylformylglycinamidine cyclo-ligase homology <PFC>
F;794-1124/Domain: phosphoribosylformylglycinamidine cyclo-ligase homology <PFCL>
F;1158-1351/Domain: phosphoribosylglycinamide formyltransferase homology <PRGF>

  Query Match             82.1%;  Score 32;  DB 1;  Length 1364;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||:|
Db        514 ELYENIG 520


RESULT 15
S45530
probable 1-phosphatidylinositol 4-kinase (EC 2.7.1.67) - yeast (Saccharomyces cerevisiae
N;Alternate names: protein L2142.4; protein YLR305c
C;Species: Saccharomyces cerevisiae
C;Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jul-2000
C;Accession: S45530; S51437
R;Yoshida, S.; Ohya, Y.; Goebl, M.; Nakano, A.; Anraku, Y.
J. Biol. Chem. 269, 1166-1172, 1994
A;Title: A novel gene, STT4, encodes a phosphatidylinositol 4-kinase in the PKC1 protein
A;Reference number: S45530; MUID:94117423; PMID:8288577
A;Accession: S45530
A;Molecule type: DNA
A;Residues: 1-1900 <YOS>
A;Cross-references: EMBL:D13717; NID:g454206; PIDN:BAA02870.1; PID:g454207
R;Pauley, A.
submitted to the EMBL Data Library, November 1994

A;Description: The sequence of S. cerevisiae cosmid L2142.
A;Reference number: S51437
A;Accession: S51437
A;Molecule type: DNA
A;Residues: 1-1900 <PAU>
A;Cross-references: EMBL:U17247; NID:g577216; PID:g577220; MIPS:YLR305c
C;Genetics:
A;Gene: SGD:STT4
A;Cross-references: SGD:S0004296; MIPS:YLR305c
A;Map position: 12R
C;Keywords: phosphotransferase; transmembrane protein
F;377-393/Domain: transmembrane #status predicted <TMM>

  Query Match             82.1%;  Score 32;  DB 2;  Length 1900;
  Best Local Similarity   71.4%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||:|
Db        141 VLYENIG 147


Search completed: July 15, 2004, 07:29:23
Job time : 12.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:20:47 ; Search time 8 Seconds
(without alignments)
58.579 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	84.6	519	1	ALGG_PSEPK	Q88nc9 pseudomonas
2	33	84.6	536	1	ALGG_PSESM	Q887q3 pseudomonas
3	33	84.6	565	1	PHR_YEAST	P05066 saccharomyc
4	33	84.6	593	1	PTNB_CHICK	Q90687 gallus gall
5	33	84.6	593	1	PTNB_HUMAN	Q06124 homo sapien
6	33	84.6	593	1	PTNB_RAT	P41499 rattus norv
7	32	82.1	178	1	YQAC_BACSU	P45900 bacillus su
8	32	82.1	1364	1	PUR2_DROPS	P16340 d trifuncti
9	32	82.1	1900	1	STT4_YEAST	P37297 saccharomyc
10	31	79.5	99	1	YLW3_CAEEL	P34406 caenorhabdi
11	31	79.5	307	1	METF_STRLI	O54235 streptomyce
12	31	79.5	357	1	RLA0_METKA	Q8tx50 methanopyru
13	31	79.5	360	1	CYSP_HEMSP	P43156 hemerocalli
14	31	79.5	367	1	FPPS_CHICK	P08836 gallus gall
15	31	79.5	468	1	HEX_ADE31	P36855 human adeno
16	31	79.5	526	1	VGLG_SIGMA	P12647 sigma virus
17	31	79.5	739	1	PPK_MYCLE	O33127 mycobacteri
18	31	79.5	742	1	PPK_MYCTU	P95111 mycobacteri
19	31	79.5	919	1	HEX_ADE12	P19900 human adeno
20	31	79.5	1302	1	MDR4_DROME	Q00449 drosophila
21	30	76.9	172	1	NUDH_VIBCH	Q9ku53 vibrio cho1
22	30	76.9	172	1	NUDH_VIBVU	Q8der5 vibrio vuln
23	30	76.9	174	1	NUDH_NEIMA	Q9jt78 neisseria m
24	30	76.9	174	1	NUDH_NEIMB	Q9jy96 neisseria m
25	30	76.9	174	1	NUDH_SHEON	Q8eh98 shewanella
26	30	76.9	174	1	NUDH_VIBPA	Q87sa4 vibrio para
27	30	76.9	205	1	VA5_VESSQ	P35786 vespula squ
28	30	76.9	228	1	GLUC_COREF	Q8rql5 corynebacte
29	30	76.9	228	1	GLUC_CORGL	P48244 corynebacte
30	30	76.9	244	1	CYAH_MYRVE	P22143 myrothecium
31	30	76.9	291	1	ENGC_STAAC	Q9kx08 staphylococ
32	30	76.9	291	1	ENGC_STAAM	Q99up7 staphylococ
33	30	76.9	342	1	ENGC_HAEDU	Q7vmf1 haemophilus
34	30	76.9	366	1	APN1_SCHPO	P50525 schizosacch
35	30	76.9	447	1	HEX_ADE04	P36850 human adeno
36	30	76.9	449	1	MURF_RICCN	Q92h60 rickettsia
37	30	76.9	525	1	YBF4_YEAST	P34219 saccharomyc
38	30	76.9	562	1	HEMA_IAJAP	P03451 influenza a
39	30	76.9	1101	1	DIA2_HUMAN	O60879 homo sapien
40	30	76.9	5255	1	BACA_BACLI	O68006 b bacitraci
41	29	74.4	94	1	RUXE_DROME	Q9v1v5 drosophila
42	29	74.4	118	1	GLBN_NOSCO	Q00812 nostoc comm
43	29	74.4	118	1	GLBN_NOSSN	P52335 nostoc sp.
44	29	74.4	170	1	NUDH_BARBA	P35640 bartonella
45	29	74.4	173	1	NUDH_RHILO	Q98f04 rhizobium l

ALIGNMENTS

RESULT 1
ALGG_PSEPK
ID ALGG_PSEPK STANDARD; PRT; 519 AA.
AC Q88NC9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Poly(beta-D-mannuronate) C5 epimerase precursor (EC 5.1.3.-).
GN ALGG OR PP1283.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
CC -!- FUNCTION: Bifunctional protein that converts poly(beta-D-
CC mannuronate) to alpha-L-guluronate and that is also part of a
CC periplasmic protein complex that serves as a scaffold that leads
CC the newly formed alginate polymer through the periplasmic space to
CC the outer membrane secretin algE (By similarity).
CC -!- PATHWAY: Alginate biosynthesis.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC -!- SIMILARITY: Belongs to the D-mannuronate C5-epimerase family.
CC -!- SIMILARITY: Contains 6 PbH1 repeats.
CC ---

CC ---
DR EMBL; AE016778; AAN66907.1; ALT_INIT.
DR TIGR; PP1283; -.
DR InterPro; IPR006633; CASH.
DR InterPro; IPR006626; PbH1.
DR SMART; SM00722; CASH; 1.
DR SMART; SM00710; PbH1; 6.
KW Alginate biosynthesis; Isomerase; Periplasmic; Repeat; Signal;
KW Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 519 POLY(BETA-D-MANNURONATE) C5 EPIMERASE.

```
FT   REPEAT      219    246       PBH1 1.
FT   REPEAT      281    303       PBH1 2.
FT   REPEAT      305    328       PBH1 3.
FT   REPEAT      330    352       PBH1 4.
FT   REPEAT      354    376       PBH1 5.
FT   REPEAT      377    399       PBH1 6.
SQ   SEQUENCE   519 AA;  57936 MW;  804D0CA87D39EDCC CRC64;

  Query Match             84.6%;  Score 33;  DB 1;  Length 519;
  Best Local Similarity   66.7%;  Pred. No. 27;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||  |:
Db        404 RLYENVAMA 412


RESULT 2
ALGG_PSESM
ID   ALGG_PSESM     STANDARD;      PRT;   536 AA.
AC   Q887Q3;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Poly(beta-D-mannuronate) C5 epimerase precursor (EC 5.1.3.-).
GN   ALGG OR PSPTO1238.
OS   Pseudomonas syringae (pv. tomato).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=323;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DC3000;
RX   MEDLINE=22834015; PubMed=12928499;
RA   Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA   Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA   Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA   Nelson W.C., Davidsen T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA   Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA   Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA   Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA   Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA   White O., Fraser C.M., Collmer A.;
RT   "The complete genome sequence of the Arabidopsis and tomato pathogen
RT   Pseudomonas syringae pv. tomato DC3000.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC   -!- FUNCTION: Bifunctional protein that converts poly(beta-D-
CC       mannuronate) to alpha-L-guluronate and that is also part of a
CC       periplasmic protein complex that serves as a scaffold that leads
CC       the newly formed alginate polymer through the periplasmic space to
CC       the outer membrane secretin algE (By similarity).
CC   -!- PATHWAY: Alginate biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the D-mannuronate C5-epimerase family.
CC   -!- SIMILARITY: Contains 5 PbH1 repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE016860; AAO54763.1; -.
DR   TIGR; PSPTO1238; -.
DR   SMART; SM00722; CASH; 1.
DR   SMART; SM00710; PbH1; 5.
KW   Alginate biosynthesis; Isomerase; Periplasmic; Repeat; Signal;
KW   Complete proteome.
FT   SIGNAL        1     36       POTENTIAL.
FT   CHAIN        37    536       POLY(BETA-D-MANNURONATE) C5 EPIMERASE.
FT   REPEAT      298    320       PBH1 1.
FT   REPEAT      322    345       PBH1 2.
FT   REPEAT      347    369       PBH1 3.
FT   REPEAT      371    393       PBH1 4.
FT   REPEAT      394    416       PBH1 5.
SQ   SEQUENCE   536 AA;  59486 MW;  B17F41A67C6AA854 CRC64;

  Query Match             84.6%;  Score 33;  DB 1;  Length 536;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|||||  |:
Db        421 KLYENVAMA 429


RESULT 3
PHR_YEAST
ID   PHR_YEAST      STANDARD;      PRT;   565 AA.
AC   P05066;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Deoxyribodipyrimidine photolyase, mitochondrial precursor
DE   (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme).
GN   PHR1 OR YOR386W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86067229; PubMed=3906569;
RA   Sancar G.B.;
RT   "Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of
RT   the PHR1 photolyase to E. coli photolyase.";
RL   Nucleic Acids Res. 13:8231-8246(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86083177; PubMed=3000886;
RA   Yasui A., Langeveld S.A.;
RT   "Homology between the photoreactivation genes of Saccharomyces
RT   cerevisiae and Escherichia coli.";
RL   Gene 36:349-355(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Delius H., Hebling U., Hofmann B.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   REVIEW.
RA   Sancar G.B., Sancar A.;
RT   "Structure and function of DNA photolyases.";
RL   Trends Biochem. Sci. 12:259-261(1987).
CC   -!- FUNCTION: This enzyme catalyzes the light-dependent monomerization
CC       (300-600 nM) of cyclobutyl pyrimidine dimers (in cis-syn
CC       configuration), which are formed between adjacent bases on the
CC       same DNA strand, upon exposure to ultraviolet radiation.
CC   -!- CATALYTIC ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine
CC       residues (in DNA).
CC   -!- COFACTOR: Contains 2 chromophores: a reduced flavin (FADH2) and a
CC       5,10-methenyltetrahydrofolate. Both chromophores are bound by non-
CC       covalent interactions.
CC   -!- SUBCELLULAR LOCATION: Nuclear and mitochondrial.
CC   -!- MISCELLANEOUS: This protein belongs to the "short wavelength-type
CC       photolyases" with an absorption maximum at about 380 nM.
CC   -!- MISCELLANEOUS: There are only 150-300 molecules of photolyase per
CC       yeast cell.
CC   -!- SIMILARITY: Belongs to the DNA photolyase class-1 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X03183; CAA26944.1; -.
DR   EMBL; M11578; AAA34875.1; -.
DR   EMBL; Z75294; CAA99718.1; -.
DR   PIR; S67298; S67298.
DR   HSSP; P00914; 1DNP.
DR   GermOnline; 143974; -.
DR   SGD; S0005913; PHR1.
DR   InterPro; IPR002081; DNA_photolyase_1.
DR   InterPro; IPR006050; DNA_photolyase_N.
DR   InterPro; IPR005101; FAD_binding_7.
DR   InterPro; IPR006051; FAD_binding_N.
DR   Pfam; PF00875; DNA_photolyase; 1.
DR   Pfam; PF03441; FAD_binding_7; 1.
DR   PRINTS; PR00147; DNAPHOTLYASE.
DR   ProDom; PD004390; FAD_binding_N; 1.
DR   PROSITE; PS00394; DNA_PHOTOLYASES_1_1; 1.
DR   PROSITE; PS00691; DNA_PHOTOLYASES_1_2; 1.
KW   Lyase; Chromophore; Flavoprotein; FAD; DNA repair; DNA-binding;
KW   Nuclear protein; Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION.
FT   CHAIN         ?    565       DEOXYRIBODIPYRIMIDINE PHOTOLYASE.
FT   DNA_BIND    421    440       H-T-H MOTIF (POTENTIAL).
FT   CONFLICT     77     77       V -> A (IN REF. 2).
FT   CONFLICT    165    165       T -> S (IN REF. 2).
FT   CONFLICT    169    169       S -> T (IN REF. 2).
FT   CONFLICT    200    200       D -> S (IN REF. 2).
FT   CONFLICT    351    351       S -> R (IN REF. 2).
FT   CONFLICT    365    365       G -> E (IN REF. 2).
FT   CONFLICT    473    473       E -> K (IN REF. 2).
SQ   SEQUENCE   565 AA;  66274 MW;  CD4FC3DA6128B97C CRC64;

  Query Match             84.6%;  Score 33;  DB 1;  Length 565;
  Best Local Similarity   55.6%;  Pred. No. 30;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||:|||::
Db         86 RLYDNVGLY 94


RESULT 4
PTNB_CHICK
ID   PTNB_CHICK     STANDARD;      PRT;   593 AA.
AC   Q90687;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein-tyrosine phosphatase, non-receptor type 11 (EC 3.1.3.48) (cSH-
DE   PTP2).
GN   PTPN11 OR SH-PTP2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythroblast;
RX   MEDLINE=97080506; PubMed=8921851;
RA   Park C.Y., LaMontagne K.R., Tonks N.K., Hayman M.J.;
RT   "Cloning and expression of the chicken protein tyrosine phosphatase
RT   SH-PTP2.";
RL   Gene 177:93-97(1996).
CC   -!- FUNCTION: This PTPase activity may directly link growth factor
CC       receptors and other signaling proteins through protein-tyrosine
CC       phosphorylation. The SH2 regions may interact with other cellular
CC       components to modulate its own phosphatase activity against
CC       interacting substrates (By similarity). May play a positive role
CC       during the stages of erythroid cell proliferation.
CC   -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC       tyrosine + phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in embryonic fibroblast,
CC       hematopoietic, erythroid, myeloid and lymphoid cells.
CC   -!- PTM: Phosphorylated by tyrosine-protein kinases (By similarity).
CC   -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC       Non-receptor class subfamily.
CC   -!- SIMILARITY: Contains 2 SH2 domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U38620; AAC60049.1; -.
DR   PIR; JC5167; JC5167.
DR   HSSP; Q06124; 2SHP.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00017; SH2; 2.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   PRINTS; PR00401; SH2DOMAIN.
DR   ProDom; PD000093; SH2; 2.
DR   SMART; SM00194; PTPc; 1.
DR   SMART; SM00252; SH2; 2.
DR   PROSITE; PS50001; SH2; 2.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase; SH2 domain; Repeat; Phosphorylation.
FT   DOMAIN        6    102       SH2 1.
FT   DOMAIN      112    216       SH2 2.
FT   DOMAIN      247    521       PROTEIN-TYROSINE PHOSPHATASE.
FT   ACT_SITE    459    459       PHOSPHOCYSTEINE INTERMEDIATE (BY
FT                                SIMILARITY).
SQ   SEQUENCE   593 AA;  67982 MW;  415231144BB43DDA CRC64;

  Query Match             84.6%;  Score 33;  DB 1;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 32;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::||||::
Db        578 RVYENVGLM 586


RESULT 5
PTNB_HUMAN
ID   PTNB_HUMAN     STANDARD;      PRT;   593 AA.
AC   Q06124;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein-tyrosine phosphatase, non-receptor type 11 (EC 3.1.3.48)
DE   (Protein-tyrosine phosphatase 2C) (PTP-2C) (PTP-1D) (SH-PTP3) (SH-
DE   PTP2) (SHP-2).
GN   PTPN11 OR PTP2C OR SHPTP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Umbilical cord;
RX   MEDLINE=93211929; PubMed=7681589;
RA   Ahmad S., Banville D.L., Zhao Z., Fischer E.H., Shen S.H.;
RT   "A widely expressed human protein-tyrosine phosphatase containing src
RT   homology 2 domains.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 90:2197-2201(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93206095; PubMed=7681217;
RA   Vogel W., Lammers R., Huang J., Ullrich A.;
RT   "Activation of a phosphotyrosine phosphatase by tyrosine
RT   phosphorylation.";
RL   Science 259:1611-1614(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-cell;
RX   MEDLINE=93106179; PubMed=1281790;
RA   Adachi M., Sekiya M., Miyachi T., Matsuno K., Hinoda Y., Imai K.,
RA   Yachi A.;
RT   "Molecular cloning of a novel protein-tyrosine phosphatase SH-PTP3
RT   with sequence similarity to the src-homology region 2.";
RL   FEBS Lett. 314:335-339(1992).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94029983; PubMed=8216283;
RA   Bastien L., Ramachandran C., Liu S., Adam M.;
RT   "Cloning, expression and mutational analysis of SH-PTP2, human
RT   protein-tyrosine phosphatase.";
RL   Biochem. Biophys. Res. Commun. 196:124-133(1993).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93087502; PubMed=1280823;
RA   Freeman R.M. Jr., Plutzky J., Neel B.G.;
RT   "Identification of a human src homology 2-containing protein-tyrosine-
RT   phosphatase: a putative homolog of Drosophila corkscrew.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:11239-11243(1992).
RN   [6]
RP   PHOSPHORYLATION BY PDGFR.
RX   MEDLINE=94316690; PubMed=8041791;
RA   Bennett A.M., Tang T.L., Sugimoto S., Walsh C.T., Neel B.G.;
RT   "Protein-tyrosine-phosphatase SHPTP2 couples platelet-derived growth
RT   factor receptor beta to Ras.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:7335-7339(1994).
RN   [7]
RP   INTERACTION WITH PTPNS1.
RX   MEDLINE=97215901; PubMed=9062191;
RA   Kharitonenkov A., Chen Z., Sures I., Wang H., Schilling J.,
RA   Ullrich A.;
RT   "A family of proteins that inhibit signalling through tyrosine kinase
RT   receptors.";
RL   Nature 386:181-186(1997).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 1-526.
RX   MEDLINE=98150850; PubMed=9491886;
RA   Hof P., Pluskey S., Dhe-Paganon S., Eck M.J., Shoelson S.E.;
RT   "Crystal structure of the tyrosine phosphatase SHP-2.";
RL   Cell 92:441-450(1998).
RN   [9]
RP   VARIANTS NS GLY-61; CYS-63; GLY-72; SER-72; ASP-76; ARG-79; VAL-282;
RP   ASP-308 AND VAL-504.
RX   MEDLINE=21583743; PubMed=11704759;
RA   Tartaglia M., Mehler E.L., Goldberg R., Zampino G., Brunner H.G.,
RA   Kremer H., van der Burgt I., Crosby A.H., Ion A., Jeffery S.,
RA   Kalidas K., Patton M.A., Kucherlapati R.S., Gelb B.D.;
RT   "Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2,
RT   cause Noonan syndrome.";
RL   Nat. Genet. 29:465-468(2001).
RN   [10]
RP   ERRATUM.
RA   Tartaglia M., Mehler E.L., Goldberg R., Zampino G., Brunner H.G.,
RA   Kremer H., van der Burgt I., Crosby A.H., Ion A., Jeffery S.,
RA   Kalidas K., Patton M.A., Kucherlapati R.S., Gelb B.D.;
RL   Nat. Genet. 29:491-491(2001).
RN   [11]
RP   ERRATUM.
RA   Tartaglia M., Mehler E.L., Goldberg R., Zampino G., Brunner H.G.,
RA   Kremer H., van der Burgt I., Crosby A.H., Ion A., Jeffery S.,
RA   Kalidas K., Patton M.A., Kucherlapati R.S., Gelb B.D.;
RL   Nat. Genet. 30:123-123(2001).
RN   [12]
RP   VARIANTS NS ALA-42; ALA-60; ASN-61; GLY-61; ASP-62; CYS-63; GLY-72;
RP   ILE-73; ASP-76; ARG-79; ALA-106; ASP-139; CYS-279; VAL-282; LEU-285;
RP   SER-285; ASP-308; SER-308; VAL-309; LYS-501 AND VAL-504, AND VARIANT
RP   NOONAN-LIKE SYNDROME SER-308.
RX   MEDLINE=21987645; PubMed=11992261;
RA   Tartaglia M., Kalidas K., Shaw A., Song X., Musat D.L.,
RA   van der Burgt I., Brunner H.G., Bertola D.R., Crosby A.H., Ion A.,
RA   Kucherlapati R.S., Jeffery S., Patton M.A., Gelb B.D.;
RT   "PTPN11 mutations in Noonan syndrome: molecular spectrum, genotype-
RT   phenotype correlation, and phenotypic heterogeneity.";
RL   Am. J. Hum. Genet. 70:1555-1563(2002).
RN   [13]
RP   VARIANTS LEOPARD SYNDROME CYS-279 AND MET-468.
RX   MEDLINE=22104852; PubMed=12058348;
RA   Digilio M.C., Conti E., Sarkozy A., Mingarelli R., Dottorini T.,
RA   Marino B., Pizzuti A., Dallapiccola B.;
RT   "Grouping of multiple-lentigines/LEOPARD and Noonan syndromes on the
RT   PTPN11 gene.";
RL   Am. J. Hum. Genet. 71:389-394(2002).
RN   [14]
RP   VARIANTS NS ASP-62; CYS-63 AND THR-502.
RX   MEDLINE=22236043; PubMed=12325025;
RA   Maheshwari M., Belmont J., Fernbach S., Ho T., Molinari L., Yakub I.,
RA   Yu F., Combes A., Towbin J., Craigen W.J., Gibbs R.;
RT   "PTPN11 mutations in Noonan syndrome type I: detection of recurrent
RT   mutations in exons 3 and 13.";
RL   Hum. Mutat. 20:298-304(2002).
RN   [15]
RP   VARIANTS NS GLY-61; CYS-63; SER-72; ILE-73; SER-285 AND ASP-308.
RX   MEDLINE=22151235; PubMed=12161469;
RA   Kosaki K., Suzuki T., Muroya K., Hasegawa T., Sato S., Matsuo N.,
RA   Kosaki R., Nagai T., Hasegawa Y., Ogata T.;
RT   "PTPN11 (protein-tyrosine phosphatase, nonreceptor-type 11) mutations
RT   in seven Japanese patients with Noonan syndrome.";
RL   J. Clin. Endocrinol. Metab. 87:3529-3533(2002).
RN   [16]
RP   VARIANTS JMML TYR-61; VAL-61; LYS-69; THR-72; VAL-72; ALA-76; GLY-76;
RP   LYS-76; VAL-76; ALA-503 AND ARG-503, VARIANTS MYELODYSPLASTIC SYNDROME
RP   VAL-60; VAL-61; LYS-69; LEU-71 AND ALA-76, VARIANTS NS ASP-62 AND
RP   ILE-73, AND VARIANT ACUTE MYELOID LEUKEMIA LYS-71.
RX   MEDLINE=22660528; PubMed=12717436;
RA   Tartaglia M., Niemeyer C.M., Fragale A., Song X., Buechner J.,
RA   Jung A., Haehlen K., Hasle H., Licht J.D., Gelb B.D.;
RT   "Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia,
RT   myelodysplastic syndromes and acute myeloid leukemia.";
RL   Nat. Genet. 34:148-150(2003).
CC   -!- FUNCTION: This PTPase activity may directly link growth factor
CC       receptors and other signaling proteins through protein-tyrosine
CC       phosphorylation. The SH2 regions may interact with other cellular
CC       components to modulate its own phosphatase activity against
CC       interacting substrates.
CC   -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC       tyrosine + phosphate.
CC   -!- SUBUNIT: Binds PTPNS1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Widely expressed; particularly abundant in
CC       heart, brain, and skeletal muscle.
CC   -!- PTM: Phosphorylation of tyrosine residues at the C-terminus by
CC       platelet-derived growth factor creates a binding site for the SH2
CC       domain of GRB2.
CC   -!- DISEASE: Defects in PTPN11 are the cause of LEOPARD syndrome
CC       [MIM:151100], an autosomal dominant disorder allelic with Noonan
CC       syndrome. The acronym LEOPARD stands for lentigines,
CC       electrocardiographic conduction abnormalities, ocular
CC       hypertelorism, pulmonic stenosis, abnormalities of genitalia,
CC       retardation of growth, and deafness.
CC   -!- DISEASE: Defects in PTPN11 are a cause of Noonan syndrome (NS)
CC       [MIM:163950]; also designated Noonan syndrome 1 (NS1). NS is an
CC       autosomal dominant disorder characterized by dysmorphic facial
CC       features, short stature, hypertelorism, cardiac anomalies,
CC       deafness, motor delay, and a bleeding diathesis. It is a
```

```
CC       genetically heterogeneous and relatively common syndrome, with an
CC       estimated incidence of 1 in 1000-2500 live births. Mutations in
CC       PTPN11 account for more than 50% of the cases. Rarely, NS is
CC       associated with juvenile myelomonocytic leukemia (JMML).
CC   -!- DISEASE: Defects in PTPN11 are a cause of Noonan-like syndrome
CC       [MIM:163955]; also known as Noonan-like/multiple giant cell lesion
CC       syndrome. It is an autosomal dominant disorder characterized by
CC       Noonan features associates with giant cell lesions of bone and
CC       soft tissue.
CC   -!- DISEASE: Defects in PTPN11 are a cause of juvenile myelomonocytic
CC       leukemia (JMML) [MIM:607785], a pediatric myelodysplastic syndrome
CC       that constitutes approximately 30% of childhood cases of
CC       myelodysplastic syndrome (MDS) and 2% of leukemia.
CC   -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC       Non-receptor class subfamily.
CC   -!- SIMILARITY: Contains 2 SH2 domains.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L08807; -; NOT_ANNOTATED_CDS.
DR   EMBL; X70766; CAA50045.1; -.
DR   EMBL; D13540; BAA02740.2; -.
DR   EMBL; L07527; AAA17022.1; -.
DR   EMBL; L03535; AAA36611.1; -.
DR   PIR; JN0805; JN0805.
DR   PDB; 2SHP; 16-FEB-99.
DR   Genew; HGNC:9644; PTPN11.
DR   MIM; 176876; -.

  Query Match             84.6%;  Score 33;  DB 1;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 32;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        578 RVYENVGLM 586


RESULT 6
PTNB_RAT
ID   PTNB_RAT       STANDARD;      PRT;   593 AA.
AC   P41499; Q62626;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein-tyrosine phosphatase, non-receptor type 11 (EC 3.1.3.48)
DE   (Protein-tyrosine phosphatase SYP).
GN   PTPN11.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=94324984; PubMed=8048963;
RA   Ding W., Zhang W.R., Sullivan K., Hashimoto N., Goldstein B.J.;
RT   "Identification of protein-tyrosine phosphatases prevalent in
RT   adipocytes by molecular cloning.";
RL   Biochem. Biophys. Res. Commun. 202:902-907(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=94216346; PubMed=7512964;
RA   Mei L., Dorherty C.A., Huganir R.L.;
RT   "RNA splicing regulates the activity of a SH2 domain-containing
RT   protein tyrosine phosphatase.";
RL   J. Biol. Chem. 269:12254-12262(1994).
RN   [3]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=93011127; PubMed=1382983;
RA   Hiraga A., Munakata H., Hata K., Suzuki Y., Tsuiki S.;
RT   "Purification and characterization of a rat liver protein-tyrosine
RT   phosphatase with sequence similarity to src-homology region 2.";
RL   Eur. J. Biochem. 209:195-206(1992).
RN   [4]
RP   PTPNS1 BINDING.
RX   MEDLINE=97215901; PubMed=9062191;
RA   Kharitonenkov A., Chen Z., Sures I., Wang H., Schilling J.,
RA   Ullrich A.;
RT   "A family of proteins that inhibit signalling through tyrosine kinase
RT   receptors.";
RL   Nature 386:181-186(1997).
CC   -!- FUNCTION: This PTPase activity may directly link growth factor
CC       receptors and other signaling proteins through protein-tyrosine
CC       phosphorylation. The SH2 regions may interact with other cellular
CC       components to modulate its own phosphatase activity against
CC       interacting substrates.
CC   -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC       tyrosine + phosphate.
CC   -!- SUBUNIT: Binds PTPNS1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- PTM: Phosphorylated by tyrosine-protein kinases (By similarity).
CC   -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC       Non-receptor class subfamily.
CC   -!- SIMILARITY: Contains 2 SH2 domains.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U09307; AAA20543.1; -.
DR   EMBL; U05963; AAA19133.1; -.
DR   PIR; A53593; A53593.
DR   HSSP; P35235; 1AYA.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00017; SH2; 2.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   PRINTS; PR00401; SH2DOMAIN.
DR   ProDom; PD000093; SH2; 2.
DR   SMART; SM00194; PTPc; 1.
DR   SMART; SM00252; SH2; 2.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
DR   PROSITE; PS50001; SH2; 2.
KW   Hydrolase; SH2 domain; Repeat; Phosphorylation.
FT   DOMAIN        6    102       SH2 1.
FT   DOMAIN      112    216       SH2 2.
FT   DOMAIN      276    517       PROTEIN-TYROSINE PHOSPHATASE.
FT   ACT_SITE    459    459       PHOSPHOCYSTEINE INTERMEDIATE (BY
FT                                SIMILARITY).
FT   CONFLICT     75     75       A -> P (IN REF. 1).
FT   CONFLICT    407    407       G -> GQALL (IN REF. 2).
FT   CONFLICT    547    547       Y -> S (IN REF. 2).
SQ   SEQUENCE   593 AA;  68033 MW;  3329F10F0F60AF48 CRC64;

  Query Match             84.6%;  Score 33;  DB 1;  Length 593;
  Best Local Similarity   55.6%;  Pred. No. 32;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
```

```
Db        578 RVYENVGLM 586


RESULT 7
YQAC_BACSU
ID   YQAC_BACSU     STANDARD;      PRT;   178 AA.
AC   P45900;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical protein yqaC precursor.
GN   YQAC OR BSU26370.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / JH642;
RX   MEDLINE=95219086; PubMed=7704261;
RA   Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
RT   "Complete nucleotide sequence of a skin element excised by DNA
RT   rearrangement during sporulation in Bacillus subtilis.";
RL   Microbiology 141:323-327(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / JH642;
RX   MEDLINE=97124195; PubMed=8969508;
RA   Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA   Kobayashi Y.;
RT   "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT   the Bacillus subtilis genome containing the skin element and many
RT   sporulation genes.";
RL   Microbiology 142:3103-3111(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [4]
RP   IDENTIFICATION.
RX   MEDLINE=96084975; PubMed=7489895;
RA   Medigue C., Moszer I., Viari A., Danchin A.;
RT   "Analysis of a Bacillus subtilis genome fragment using a co-operative
RT   computer system prototype.";
RL   Gene 165:GC37-GC51(1995).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D32216; BAA06916.1; -.
DR   EMBL; D84432; BAA12377.1; -.
DR   EMBL; Z99117; CAB14578.1; -.
DR   PIR; B69944; B69944.
DR   SubtiList; BG11254; yqaC.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    178       HYPOTHETICAL PROTEIN YQAC.
SQ   SEQUENCE   178 AA;  20702 MW;  DD2DE09D65CF882E CRC64;

  Query Match             82.1%;  Score 32;  DB 1;  Length 178;
  Best Local Similarity   55.6%;  Pred. No. 14;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||:| ||:
Db        129 SLYDNAGME 137


RESULT 8
PUR2_DROPS
ID   PUR2_DROPS     STANDARD;      PRT;  1364 AA.
AC   P16340;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Trifunctional purine biosynthetic protein adenosine-3 [Includes:
DE   Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide
DE   ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase);
DE   Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1) (AIRS)
DE   (Phosphoribosyl-aminoimidazole synthetase) (AIR synthase);
DE   Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2) (GART) (GAR
DE   transformylase) (5'-phosphoribosylglycinamide transformylase)].
GN   ADE3 OR GART.
OS   Drosophila pseudoobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7237;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=EST10;
RX   MEDLINE=88112752; PubMed=3123310;
RA   Henikoff S., Eghtedarzadeh M.K.;
RT   "Conserved arrangement of nested genes at the Drosophila Gart locus.";
RL   Genetics 117:711-725(1987).
CC   -!- CATALYTIC ACTIVITY: ATP + 5-phospho-D-ribosylamine + glycine = ADP
CC       + phosphate + N(1)-(5-phospho-D-ribosyl)glycinamide.
CC   -!- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + N(1)-(5-phospho-D-
CC       ribosyl)glycinamide = tetrahydrofolate + N(2)-formyl-N(1)-(5-
CC       phospho-D-ribosyl)glycinamide.
CC   -!- CATALYTIC ACTIVITY: ATP + 2-(formamido)-N(1)-(5-phospho-D-
CC       ribosyl)acetamidine = ADP + phosphate + 5-amino-1-(5-phospho-D-
CC       ribosyl)imidazole.
CC   -!- PATHWAY: De novo purine biosynthesis; second step.
CC   -!- PATHWAY: De novo purine biosynthesis; third step.
CC   -!- PATHWAY: De novo purine biosynthesis; fifth step.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P16340-1; Sequence=Displayed;
CC       Name=Short;
```

```
CC         IsoId=P16340-2; Sequence=VSP_005514, VSP_005515;
CC   -!- SIMILARITY: In the N-terminal section; belongs to the GARS family.
CC   -!- SIMILARITY: In the central section; belongs to the AIR synthase
CC       family.
CC   -!- SIMILARITY: TO OTHER AIRS AND GART FROM BACTERIA AND EUKARYOTES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X06285; CAA29611.1; -.
DR   PIR; S01204; AJFFPP.
DR   HSSP; P08179; 1GAR.
DR   FlyBase; FBgn0012717; Dpse\ade3.
DR   InterPro; IPR000728; AIR_synth.
DR   InterPro; IPR002376; formyl_transf.
DR   InterPro; IPR000115; Gars.
DR   InterPro; IPR001555; GART_AS.
DR   InterPro; IPR004733; PurM_cligase.
DR   InterPro; IPR004607; PurN.
DR   Pfam; PF00586; AIRS; 2.
DR   Pfam; PF02769; AIRS_C; 2.
DR   Pfam; PF00551; formyl_transf; 1.
DR   Pfam; PF01071; GARS; 1.
DR   Pfam; PF02842; GARS_B; 1.
DR   Pfam; PF02843; GARS_C; 1.
DR   Pfam; PF02844; GARS_N; 1.
DR   TIGRFAMs; TIGR00877; purD; 1.
DR   TIGRFAMs; TIGR00878; purM; 2.
DR   TIGRFAMs; TIGR00639; PurN; 1.
DR   PROSITE; PS00184; GARS; 1.
DR   PROSITE; PS00373; GART; 1.
KW   Multifunctional enzyme; Purine biosynthesis; Ligase; Transferase;
KW   Alternative splicing.
FT   DOMAIN        1    434       GARS.
FT   DOMAIN      435   1154       AIRS.
FT   DOMAIN     1155   1364       GART.
FT   ACT_SITE   1301   1301       BY SIMILARITY.
FT   VARSPLIC    434    434       I -> M (in isoform Short).
FT                                /FTId=VSP_005514.
FT   VARSPLIC    435   1364       Missing (in isoform Short).
FT                                /FTId=VSP_005515.
SQ   SEQUENCE   1364 AA;  145693 MW;  BBD4B5166FF4D301 CRC64;

  Query Match             82.1%;  Score 32;  DB 1;  Length 1364;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||:|
Db        514 ELYENIG 520


RESULT 9
STT4_YEAST
ID   STT4_YEAST     STANDARD;      PRT;  1900 AA.
AC   P37297;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Phosphatidylinositol 4-kinase STT4 (EC 2.7.1.67) (PI4-kinase)
DE   (PtdIns-4-kinase).
GN   STT4 OR YLR305C OR L2142.4.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94117423; PubMed=8288577;
RA   Yoshida S., Goebl M., Ohya Y., Nakano A., Anraku Y.;
RT   "A novel gene, STT4, encodes a phosphatidylinositol 4-kinase in the
RT   PKC1 protein kinase pathway of Saccharomyces cerevisiae.";
RL   J. Biol. Chem. 269:1166-1172(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313267; PubMed=9169871;
RA   Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA   Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA   Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA   Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA   Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA   Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA   Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA   Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA   Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA   Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA   Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL   Nature 387:87-90(1997).
CC   -!- FUNCTION: Acts on phosphatidylinositol (PI) in the first
CC       committed step in the production of the second messenger
CC       inositol-1,4,5,-trisphosphate. STT4 functions in PKC1 protein
CC       kinase pathway.
CC   -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC       1-phosphatidyl-1D-myo-inositol 4-phosphate.
CC   -!- SIMILARITY: Belongs to the PI3/PI4-kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D13717; BAA02870.1; -.
DR   EMBL; U17247; AAB67358.1; -.
DR   EMBL; U17243; AAB67354.1; -.
DR   PIR; S45530; S45530.
DR   GermOnline; 142368; -.
DR   SGD; S0004296; STT4.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0004430; F:1-phosphatidylinositol 4-kinase activity; IMP.
DR   GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; IMP.
DR   GO; GO:0000165; P:MAPKKK cascade; IDA.
DR   GO; GO:0006646; P:phosphatidylethanolamine biosynthesis; IMP.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   InterPro; IPR001263; PI3Ka.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   Pfam; PF00613; PI3Ka; 1.
DR   SMART; SM00145; PI3Ka; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Kinase.
FT   DOMAIN     1643   1882       PI3K/PI4K.
SQ   SEQUENCE   1900 AA;  214605 MW;  F210BAF987BA276A CRC64;

  Query Match             82.1%;  Score 32;  DB 1;  Length 1900;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||:|
Db        141 VLYENIG 147
```

```
RESULT 10
YLW3_CAEEL
ID   YLW3_CAEEL     STANDARD;      PRT;    99 AA.
AC   P34406;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein F22B7.3 in chromosome III.
GN   F22B7.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L12018; AAA65463.1; -.
DR   PIR; S44632; S44632.
DR   WormPep; F22B7.3; CE00156.
KW   Hypothetical protein.
SQ   SEQUENCE   99 AA;  11665 MW;  78FC94DBD3C8B585 CRC64;

  Query Match             79.5%;  Score 31;  DB 1;  Length 99;
  Best Local Similarity   71.4%;  Pred. No. 13;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YENVGMX 9
              |||:||:
Db         21 YENLGMF 27


RESULT 11
METF_STRLI
ID   METF_STRLI     STANDARD;      PRT;   307 AA.
AC   O54235;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN   METF.
OS   Streptomyces lividans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1916;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=66 / 1326;
RX   MEDLINE=98175688; PubMed=9515933;
RA   Blanco J., Coque J.R., Martin J.;
RT   "The folate branch of the methionine biosynthesis pathway in
RT   Streptomyces lividans: disruption of the 5,10-
RT   methylenetetrahydrofolate reductase gene leads to methionine
RT   auxotrophy.";
RL   J. Bacteriol. 180:1586-1591(1998).
CC   -!- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + acceptor = 5,10-
CC       methylenetetrahydrofolate + reduced acceptor.
CC   -!- COFACTOR: FAD (By similarity).
CC   -!- PATHWAY: Methionine biosynthesis.
CC   -!- SIMILARITY: Belongs to the methylenetetrahydrofolate reductase
CC       (EC 1.5.1.20/EC 1.7.99.5) family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001630; CAA04885.1; -.
DR   HSSP; P00394; 1B5T.
DR   InterPro; IPR004620; Fadh2_bact.
DR   InterPro; IPR003171; Mehydrof_redctse.
DR   Pfam; PF02219; MTHFR; 1.
DR   TIGRFAMs; TIGR00676; fadh2; 1.
KW   Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis.
SQ   SEQUENCE   307 AA;  33267 MW;  0CA09C336036D8A9 CRC64;

  Query Match             79.5%;  Score 31;  DB 1;  Length 307;
  Best Local Similarity   44.4%;  Pred. No. 42;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
                ::|||:|::
Db        294 EIYENLGLH 302


RESULT 12
RLA0_METKA
ID   RLA0_METKA     STANDARD;      PRT;   357 AA.
AC   Q8TX50;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Acidic ribosomal protein P0 homolog (L10E).
GN   RPLP0 OR MK0826.
OS   Methanopyrus kandleri.
OC   Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC   Methanopyrus.
OX   NCBI_TaxID=2320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AV19 / DSM 6324 / JCM 9639;
RX   MEDLINE=21927647; PubMed=11930014;
RA   Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA   Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA   Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA   Malykh A.G., Koonin E.V., Kozyavkin S.A.;
RT   "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT   and monophyly of archaeal methanogens.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC   -!- FUNCTION: Ribosomal protein P0 is the functional equivalent of
CC       E.coli protein L10.
CC   -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; AE010373; AAM02039.1; ALT_INIT.
DR   HAMAP; MF_00280; -; 1.
DR   InterPro; IPR001790; Ribosomal_L10.
DR   Pfam; PF00466; Ribosomal_L10; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   357 AA;  39250 MW;  470294320ADBBE5C CRC64;

  Query Match             79.5%;  Score 31;  DB 1;  Length 357;
  Best Local Similarity   71.4%;  Pred. No. 49;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YENVGMX 9
              |||||::
Db         35 YENVGLV 41


RESULT 13
CYSP_HEMSP
ID   CYSP_HEMSP     STANDARD;      PRT;   360 AA.
AC   P43156;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Thiol protease SEN102 precursor (EC 3.4.22.-).
GN   SEN102.
OS   Hemerocallis sp. (Daylily).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
OC   Hemerocallidaceae; Hemerocallis.
OX   NCBI_TaxID=29711;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Craddle Song; TISSUE=Petal;
RX   MEDLINE=95359413; PubMed=7632925;
RA   Valpuesta V., Lange N., Guerrero C., Reid M.;
RT   "Up-regulation of a cysteine protease accompanies the ethylene-
RT   insensitive senescence of daylily (Hemerocallis) flowers.";
RL   Plant Mol. Biol. 28:575-582(1995).
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
CC   -!- SIMILARITY: Belongs to peptidase family C1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74406; CAA52425.1; -.
DR   PIR; S57777; S57777.
DR   HSSP; P07711; 1CJL.
DR   MEROPS; C01.010; -.
DR   InterPro; IPR000886; ER_target_S.
DR   InterPro; IPR000668; Peptidase_C1.
DR   InterPro; IPR000169; SHprot_acsite.
DR   Pfam; PF00112; Peptidase_C1; 1.
DR   PRINTS; PR00705; PAPAIN.
DR   ProDom; PD000158; Peptidase_C1; 1.
DR   SMART; SM00645; Pept_C1; 1.
DR   PROSITE; PS00014; ER_TARGET; 1.
DR   PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR   PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
KW   Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal;
KW   Endoplasmic reticulum.
FT   SIGNAL        1     20       POTENTIAL.
FT   PROPEP       21    133       ACTIVATION PEPTIDE (POTENTIAL).
FT   CHAIN       134    360       THIOL PROTEASE SEN102.
FT   ACT_SITE    154    154       BY SIMILARITY.
FT   ACT_SITE    289    289       BY SIMILARITY.
FT   ACT_SITE    310    310       BY SIMILARITY.
FT   CARBOHYD    353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   SITE        357    360       PREVENT SECRETION FROM ER (POTENTIAL).
SQ   SEQUENCE   360 AA;  39914 MW;  808A3D252D2A2C63 CRC64;

  Query Match             79.5%;  Score 31;  DB 1;  Length 360;
  Best Local Similarity   71.4%;  Pred. No. 50;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
                ::|||||
Db        121 FMYENVG 127


RESULT 14
FPPS_CHICK
ID   FPPS_CHICK     STANDARD;      PRT;   367 AA.
AC   P08836;
DT   01-NOV-1988 (Rel. 09, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Farnesyl pyrophosphate synthetase (FPP synthetase) (FPS) (Farnesyl
DE   diphosphate synthetase) [Includes: Dimethylallyltransferase
DE   (EC 2.5.1.1); Geranyltranstransferase (EC 2.5.1.10)].
GN   FDPS.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE OF 187-216.
RC   TISSUE=Liver;
RX   MEDLINE=82000466; PubMed=7272273;
RA   Brems D.N., Bruenger E., Rillings H.C.;
RT   "Isolation and characterization of a photoaffinity-labeled peptide
RT   from the catalytic site of prenyltransferase.";
RL   Biochemistry 20:3711-3718(1981).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 20-367.
RC   TISSUE=Liver;
RX   MEDLINE=94368786; PubMed=8086404;
RA   Tarshis L.C., Yan M., Poulter C.D., Sacchettini J.C.;
RT   "Crystal structure of recombinant farnesyl diphosphate synthase at
RT   2.6-A resolution.";
RL   Biochemistry 33:10871-10877(1994).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-367.
RC   TISSUE=Liver;
RX   MEDLINE=97140274; PubMed=8986756;
RA   Tarshis L.C., Proteau P.J., Kellogg B.A., Sacchettini J.C.,
RA   Poulter C.D.;
RT   "Regulation of product chain length by isoprenyl diphosphate
RT   synthases.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:15018-15023(1996).
CC   -!- FUNCTION: Catalyzes the sequential condensation of isopentenyl
CC       pyrophosphate with the allylic pyrophosphates, dimethylallyl
CC       pyrophosphate, and then with the resultant geranylpyrophosphate to
CC       the ultimate product farnesyl pyrophosphate.
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- PATHWAY: Isoprene biosynthesis, cholesterol biosynthesis.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the FPP/GGPP synthetase family.
DR   PDB; 1FPS; 10-JUL-95.
DR   PDB; 1UBV; 12-MAR-97.
DR   PDB; 1UBW; 12-MAR-97.
DR   PDB; 1UBX; 12-MAR-97.
DR   PDB; 1UBY; 12-MAR-97.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   InterPro; IPR008949; Terpenoid_synth.
```

```
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Isoprene biosynthesis; Cholesterol biosynthesis;
KW   3D-structure.
FT   ACT_SITE    206    206
FT   CONFLICT    191    191       T -> G (IN REF. 1).
FT   CONFLICT    201    203       HFS -> TFQ (IN REF. 1).
FT   CONFLICT    210    212       IVK -> FVP (IN REF. 1).
FT   CONFLICT    215    216       TA -> AM (IN REF. 1).
FT   HELIX        21     33
FT   TURN         34     34
FT   HELIX        35     43
FT   TURN         44     46
FT   TURN         50     52
FT   HELIX        53     66
FT   HELIX        73     85
FT   HELIX        88     90
FT   HELIX        93    121
FT   TURN        122    122
FT   STRAND      125    126
FT   TURN        127    128
FT   STRAND      129    130
FT   HELIX       132    134
FT   TURN        136    140
FT   HELIX       141    161
FT   TURN        162    163
FT   TURN        165    166
FT   HELIX       167    191
FT   TURN        194    195
FT   TURN        200    201
FT   HELIX       204    214
FT   TURN        215    215
FT   HELIX       216    219
FT   TURN        220    220
FT   HELIX       221    231
FT   TURN        232    232
FT   HELIX       236    263
FT   TURN        266    267
FT   HELIX       283    291
FT   HELIX       294    303
FT   TURN        304    305
FT   HELIX       309    322
FT   TURN        323    323
FT   HELIX       324    346
FT   TURN        352    352
FT   HELIX       353    362
SQ   SEQUENCE   367 AA;  42153 MW;  BB23D29D62CD842B CRC64;

  Query Match             79.5%;  Score 31;  DB 1;  Length 367;
  Best Local Similarity   66.7%;  Pred. No. 51;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||| |||:
Db        317 ELYEAVGMR 325


RESULT 15
HEX_ADE31
ID   HEX_ADE31      STANDARD;      PRT;   468 AA.
AC   P36855;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Hexon protein (Late protein 2) (Fragment).
GN   PII.
OS   Human adenovirus type 31.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=10529;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VRL 15/62;
RX   MEDLINE=94294642; PubMed=8023012;
RA   Pring-Akerblom P., Adrian T.;
RT   "Type- and group-specific polymerase chain reaction for adenovirus
RT   detection.";
RL   Res. Virol. 145:25-35(1994).
CC   -!- FUNCTION: This protein is one of the structural proteins in the
CC       viral coat and is synthesized during late infection.
CC   -!- SUBUNIT: Homotrimer (By similarity).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74661; CAA52725.1; -.
DR   PIR; S37217; S37217.
DR   HSSP; P03277; 1DHX.
DR   InterPro; IPR000736; Adeno_hexon.
DR   Pfam; PF01065; Adeno_hexon; 1.
DR   ProDom; PD002815; Adeno_hexon; 1.
KW   Coat protein; Hexon protein; Late protein.
FT   NON_TER       1      1
FT   NON_TER     468    468
SQ   SEQUENCE   468 AA;  52100 MW;  8727BFA49179CE68 CRC64;

  Query Match             79.5%;  Score 31;  DB 1;  Length 468;
  Best Local Similarity   55.6%;  Pred. No. 66;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|| |||::
Db        341 FLYSNVGLY 349


Search completed: July 15, 2004, 07:27:02
Job time : 9 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 07:25:27 ; Search time 33 Seconds
(without alignments)
86.050 Million cell updates/sec

Title: SEQ1MOD
Perfect score: 39
Sequence: 1 XLYENVGMX 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	87.2	342	13	Q9IB95	Q9ib95 potamotrygo
2	34	87.2	430	16	Q9RIV4	Q9riv4 streptomyce
3	34	87.2	688	5	Q9TYZ8	Q9tyz8 caenorhabdi
4	33	84.6	56	16	Q89RP9	Q89rp9 bradyrhizob
5	33	84.6	594	13	Q7ZW17	Q7zw17 brachydanio
6	33	84.6	595	13	Q92124	Q92124 xenopus lae
7	33	84.6	597	11	Q64509	Q64509 mus musculu
8	33	84.6	613	16	Q89P36	Q89p36 bradyrhizob
9	33	84.6	700	5	P90929	P90929 caenorhabdi
10	33	84.6	1353	5	Q9VM53	Q9vm53 drosophila
11	32	82.1	133	16	Q88GB4	Q88gb4 pseudomonas
12	32	82.1	231	10	Q9SV79	Q9sv79 arabidopsis
13	32	82.1	283	8	Q8SML6	Q8sml6 dunaliella
14	32	82.1	352	16	Q9X169	Q9x169 thermotoga
15	32	82.1	365	16	Q8RBB7	Q8rbb7 thermoanaer
16	32	82.1	367	16	Q927F5	Q927f5 listeria in
17	32	82.1	369	16	Q8Y3Z2	Q8y3z2 listeria mo
18	32	82.1	403	16	Q88VH6	Q88vh6 lactobacill
19	32	82.1	453	16	Q8DVF8	Q8dvf8 streptococc
20	32	82.1	610	16	Q88EC4	Q88ec4 pseudomonas
21	32	82.1	617	16	Q97HG4	Q97hg4 clostridium
22	32	82.1	783	16	Q839N9	Q839n9 enterococcu
23	32	82.1	1307	5	Q8MT77	Q8mt77 drosophila
24	32	82.1	1817	13	Q7SZF6	Q7szf6 xenopus lae
25	31	79.5	102	16	Q82VV7	Q82vv7 nitrosomona
26	31	79.5	135	11	Q7TN07	Q7tn07 mus musculu
27	31	79.5	149	16	O33283	O33283 mycobacteri
28	31	79.5	149	16	Q7TY00	Q7ty00 mycobacteri
29	31	79.5	165	16	Q7VH04	Q7vh04 helicobacte
30	31	79.5	166	5	O01517	O01517 caenorhabdi
31	31	79.5	169	2	Q9XDA0	Q9xda0 clostridium
32	31	79.5	183	11	Q8CES0	Q8ces0 mus musculu
33	31	79.5	219	17	Q9V1X0	Q9v1x0 pyrococcus
34	31	79.5	221	10	Q8W437	Q8w437 vigna radia
35	31	79.5	282	16	Q83E64	Q83e64 coxiella bu
36	31	79.5	307	16	Q9S2V3	Q9s2v3 streptomyce
37	31	79.5	307	16	Q82AF8	Q82af8 streptomyce
38	31	79.5	321	16	Q8CQB9	Q8cqb9 staphylococ
39	31	79.5	356	16	Q97EY7	Q97ey7 clostridium
40	31	79.5	377	5	Q95RC0	Q95rc0 drosophila
41	31	79.5	383	5	Q9W5A3	Q9w5a3 drosophila
42	31	79.5	401	16	Q88UV0	Q88uv0 lactobacill
43	31	79.5	402	5	O46309	O46309 drosophila
44	31	79.5	406	16	Q9CDQ4	Q9cdq4 lactococcus
45	31	79.5	429	16	Q834E7	Q834e7 enterococcu

ALIGNMENTS

RESULT 1
Q9IB95
ID Q9IB95 PRELIMINARY; PRT; 342 AA.
AC Q9IB95;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE RyPTPN6b protein (Fragment).
GN RYPTPN6B.
OS Potamotrygon motoro (South American freshwater stingray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC Myliobatiformes; Myliobatoidei; Potamotrygonidae; Potamotrygon.
OX NCBI_TaxID=86373;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219325; PubMed=10754074;
RA Ono-Koyanagi K., Suga H., Katoh K., Miyata T.;
RT "Protein tyrosine phosphatases from amphioxus, hagfish, and ray:
RT divergence of tissue-specific isoform genes in the early evolution of
RT vertebrates.";
RL J. Mol. Evol. 50:302-311(2000).
DR EMBL; AB033591; BAA95198.1; -.
DR HSSP; Q06124; 2SHP.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT NON_TER 1 1
SQ SEQUENCE 342 AA; 39532 MW; FCAEEA69442A4677 CRC64;

```
  Query Match             87.2%;  Score 34;  DB 13;  Length 342;
  Best Local Similarity   55.6%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        327 RIYENVGLM 335


RESULT 2
Q9RIV4
ID   Q9RIV4      PRELIMINARY;      PRT;   430 AA.
AC   Q9RIV4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative solute-binding protein.
GN   SCO0952 OR SCM11.07C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Oliver K., Harris D.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939107; CAB61918.1; -.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR006059; SBP_bac_1.
DR   Pfam; PF01547; SBP_bac_1; 1.
KW   Complete proteome.
SQ   SEQUENCE   430 AA;  46312 MW;  151F92EBF5B9C754 CRC64;

  Query Match             87.2%;  Score 34;  DB 16;  Length 430;
  Best Local Similarity   55.6%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||||:|::
Db        288 NLYENIGIT 296


RESULT 3
Q9TYZ8
ID   Q9TYZ8      PRELIMINARY;      PRT;   688 AA.
AC   Q9TYZ8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   F58E2.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Goela D., Delehaunty A.;
RT   "The sequence of C. elegans cosmid F58E2.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF100659; AAC68967.1; -.
DR   PIR; T33708; T33708.
DR   WormPep; F58E2.4; CE17132.
DR   InterPro; IPR002900; DUF38.
DR   InterPro; IPR001810; F-box.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF01827; FTH; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   688 AA;  79592 MW;  338530655E757124 CRC64;

  Query Match             87.2%;  Score 34;  DB 5;  Length 688;
  Best Local Similarity   55.6%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        393 LIYENVGLS 401


RESULT 4
Q89RP9
ID   Q89RP9      PRELIMINARY;      PRT;    56 AA.
AC   Q89RP9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Bsl2713 protein.
GN   BSL2713.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
```

```
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005944; BAC47978.1; -.
KW   Complete proteome.
SQ   SEQUENCE   56 AA;  6170 MW;  C225ADEBA8E90BA1 CRC64;

  Query Match             84.6%;  Score 33;  DB 16;  Length 56;
  Best Local Similarity   55.6%;  Pred. No. 28;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :|:||:||:
Db         43 SLHENIGMK 51


RESULT 5
Q7ZW17
ID   Q7ZW17      PRELIMINARY;      PRT;   594 AA.
AC   Q7ZW17;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to protein tyrosine phosphatase, non-receptor type 11 (Noonan
DE   syndrome 1).
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045328; AAH45328.1; -.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR003595; PTPc_motif.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00017; SH2; 2.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   PRINTS; PR00401; SH2DOMAIN.
DR   ProDom; PD000093; SH2; 2.
DR   SMART; SM00194; PTPc; 1.
DR   SMART; SM00404; PTPc_motif; 1.
DR   SMART; SM00252; SH2; 2.
DR   PROSITE; PS50001; SH2; 2.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Receptor.
SQ   SEQUENCE   594 AA;  67653 MW;  68DDBCA9B93BE2F8 CRC64;

  Query Match             84.6%;  Score 33;  DB 13;  Length 594;
  Best Local Similarity   55.6%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        579 RVYENVGLM 587


RESULT 6
Q92124
ID   Q92124      PRELIMINARY;      PRT;   595 AA.
AC   Q92124;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Phosphotyrosyl-protein phosphatase (EC 3.1.3.48).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocyte;
RX   MEDLINE=95163101; PubMed=7859288;
RA   Tang T.L., Freeman R.M. Jr., O'Reilly A.M., Neel B.G., Sokol S.Y.;
RT   "The SH2-containing protein-tyrosine phosphatase SH-PTP2 is required
RT   upstream of MAP kinase for early Xenopus development.";
RL   Cell 80:473-483(1995).
DR   EMBL; U15287; AAA65731.1; -.
DR   PIR; A55651; A55651.
DR   HSSP; P35235; 1AYA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004727; F:prenylated protein tyrosine phosphatase act. . .; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00017; SH2; 2.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   PRINTS; PR00401; SH2DOMAIN.
DR   ProDom; PD000093; SH2; 2.
DR   SMART; SM00194; PTPc; 1.
DR   SMART; SM00252; SH2; 2.
DR   PROSITE; PS50001; SH2; 2.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
SQ   SEQUENCE   595 AA;  68249 MW;  4F39FCF1E8F8D726 CRC64;

  Query Match             84.6%;  Score 33;  DB 13;  Length 595;
  Best Local Similarity   55.6%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        580 RVYENVGLL 588


RESULT 7
Q64509
ID   Q64509      PRELIMINARY;      PRT;   597 AA.
AC   Q64509;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein tyrosine phosphatase (EC 3.1.3.48).
GN   PTPN11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Balb/c; TISSUE=Brain;
RX   MEDLINE=96411777; PubMed=8810330;
RA   Ohnishi H., Kubota M., Ohtake A., Sato K., Sano S.;
RT   "Activation of protein-tyrosine phosphatase SH-PTP2 by a tryrosine-
RT   based activation motif of a novel brain molecule.";
RL   J. Biol. Chem. 271:25569-25574(1996).
DR   EMBL; D84372; BAA12328.1; -.
DR   HSSP; P35235; 1AYA.
```

```
DR   MGD; MGI:99511; Ptpn11.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0007409; P:axonogenesis; IMP.
DR   GO; GO:0048011; P:NGF receptor signaling pathway; IMP.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00017; SH2; 2.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   PRINTS; PR00401; SH2DOMAIN.
DR   ProDom; PD000093; SH2; 2.
DR   SMART; SM00194; PTPc; 1.
DR   SMART; SM00252; SH2; 2.
DR   PROSITE; PS50001; SH2; 2.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
SQ   SEQUENCE   597 AA;  68460 MW;  C742BED37E39EA23 CRC64;

  Query Match             84.6%;  Score 33;  DB 11;  Length 597;
  Best Local Similarity   55.6%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              ::|||||::
Db        582 RVYENVGLM 590


RESULT 8
Q89P36
ID   Q89P36      PRELIMINARY;      PRT;   613 AA.
AC   Q89P36;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Bll3647 protein.
GN   BLL3647.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005948; BAC48912.1; -.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR001296; Glyco_trans_1.
DR   Pfam; PF00534; Glycos_transf_1; 1.
KW   Complete proteome.
SQ   SEQUENCE   613 AA;  68932 MW;  53226C6AD8B83AE1 CRC64;

  Query Match             84.6%;  Score 33;  DB 16;  Length 613;
  Best Local Similarity   85.7%;  Pred. No. 3.3e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||||
Db        317 SLYENVG 323


RESULT 9
P90929
ID   P90929      PRELIMINARY;      PRT;   700 AA.
AC   P90929; O02514; Q19156; Q93441;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   F07C6.4b protein.
GN   F07C6.4 OR F07C6.4B.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lightning J.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Steward C.A.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z81102; CAB03204.1; -.
DR   EMBL; Z69659; CAB03204.1; JOINED.
DR   EMBL; Z69659; CAA93486.1; -.
DR   EMBL; Z81102; CAA93486.1; JOINED.
DR   PIR; T20550; T20550.
DR   WormPep; F07C6.4b; CE18569.
KW   Hypothetical protein.
SQ   SEQUENCE   700 AA;  77598 MW;  293869E242E3C6DA CRC64;

  Query Match             84.6%;  Score 33;  DB 5;  Length 700;
  Best Local Similarity   85.7%;  Pred. No. 3.8e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||||
Db        398 PLYENVG 404


RESULT 10
Q9VM53
ID   Q9VM53      PRELIMINARY;      PRT;  1353 AA.
AC   Q9VM53;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG31628 protein.
GN   CG31908 OR CG8761.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
```

```
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003615; AAF52474.2; -.
DR   HSSP; P08179; 1GAR.
DR   FlyBase; FBgn0000053; ade3.
DR   FlyBase; FBgn0051908; CG31908.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0004637; F:phosphoribosylamine-glycine ligase activity; IEA.
DR   GO; GO:0004641; F:phosphoribosylformylglycinamidine cyclo-lig. . .; IEA.
DR   GO; GO:0004644; F:phosphoribosylglycinamide formyltransferase. . .; IEA.
DR   GO; GO:0006189; P:'de novo' IMP biosynthesis; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   GO; GO:0009113; P:purine base biosynthesis; IEA.
DR   InterPro; IPR000728; AIR_synth.
DR   InterPro; IPR002376; formyl_transf.
DR   InterPro; IPR000115; Gars.
DR   InterPro; IPR001555; GART_AS.
DR   InterPro; IPR004733; PurM_cligase.
DR   InterPro; IPR004607; PurN.
DR   Pfam; PF00586; AIRS; 2.
DR   Pfam; PF02769; AIRS_C; 2.
DR   Pfam; PF00551; formyl_transf; 1.
DR   Pfam; PF01071; GARS; 1.
DR   Pfam; PF02842; GARS_B; 1.
DR   Pfam; PF02843; GARS_C; 1.
DR   Pfam; PF02844; GARS_N; 1.
DR   TIGRFAMs; TIGR00877; purD; 1.
DR   TIGRFAMs; TIGR00878; purM; 2.
DR   TIGRFAMs; TIGR00639; PurN; 1.
DR   PROSITE; PS00184; GARS; 1.
DR   PROSITE; PS00373; GART; 1.
SQ   SEQUENCE   1353 AA;  144525 MW;  3F193005CF1D7ACB CRC64;

  Query Match           84.6%;  Score 33;  DB 5;  Length 1353;
  Best Local Similarity   85.7%;  Pred. No. 7.5e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||||
Db        514 ELYENVG 520


RESULT 11
Q88GB4
ID   Q88GB4      PRELIMINARY;      PRT;   133 AA.
AC   Q88GB4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Conserved hypothetical protein.
GN   PP3810.
OS   Pseudomonas putida (strain KT2440).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=160488;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22423060; PubMed=12534463;
RA   Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA   Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA   Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA   Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA   Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA   Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA   Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA   Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative analysis of the
RT   metabolically versatile Pseudomonas putida KT2440.";
RL   Environ. Microbiol. 4:799-808(2002).
DR   EMBL; AE016788; AAN69404.1; -.
DR   TIGR; PP3810; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   133 AA;  15335 MW;  82D775532F236679 CRC64;

  Query Match           82.1%;  Score 32;  DB 16;  Length 133;
  Best Local Similarity   55.6%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||||:|::
Db         63 RLYENLGIR 71


RESULT 12
```

```
Q9SV79
ID   Q9SV79      PRELIMINARY;      PRT;   231 AA.
AC   Q9SV79;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   F25G13.3 OR AT4G12900.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
RA   Mayer K.F.X., Lemcke K., Schueller C.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-13 FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Robben J., Grymonprez B., Volckaert G, Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL079349; CAB53090.1; -.
DR   EMBL; AL161535; CAB78332.1; -.
DR   PIR; H85138; H85138.
DR   GO; GO:0004182; F:carboxypeptidase A activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR004911; GILT.
DR   InterPro; IPR000834; Peptidase_M14.
DR   Pfam; PF03227; GILT; 1.
DR   PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   231 AA;  26025 MW;  734109A78E942295 CRC64;

  Query Match             82.1%;  Score 32;  DB 10;  Length 231;
  Best Local Similarity   71.4%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVG 7
              :||||:|
Db        182 PLYENIG 188


RESULT 13
Q8SML6
ID   Q8SML6      PRELIMINARY;      PRT;   283 AA.
AC   Q8SML6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chloroplast large-subunit ribosomal RNA (rrnL), site-specific DNA
DE   endonuclease I-DpaI genes.
OS   Dunaliella parva.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Dunaliellaceae; Dunaliella.
OX   NCBI_TaxID=3048;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turmel M., Otis C., Mercier J.-P., Gutell R.R., Lemieux C.;
RT   "Distribution of group I introns in the chloroplast large subunit rRNA
RT   gene of green algae.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L43540; AAL77562.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0004519; F:endonuclease activity; IEA.
DR   InterPro; IPR004860; LAGLIDADG_2.
DR   Pfam; PF03161; LAGLIDADG_2; 1.
KW   Chloroplast.
SQ   SEQUENCE   283 AA;  32080 MW;  2BD710EC7BBC1E82 CRC64;

  Query Match             82.1%;  Score 32;  DB 8;  Length 283;
  Best Local Similarity   55.6%;  Pred. No. 2.4e+02;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||||:|::
Db        197 ALYENLGIE 205


RESULT 14
Q9X169
ID   Q9X169      PRELIMINARY;      PRT;   352 AA.
AC   Q9X169;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein TM1348.
GN   TM1348.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
DR   EMBL; AE001789; AAD36419.1; -.
DR   PIR; D72264; D72264.
DR   TIGR; TM1348; -.
DR   InterPro; IPR001440; TPR.
DR   InterPro; IPR008941; TPR-like.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   352 AA;  41237 MW;  47EF0B432D421CB8 CRC64;

  Query Match             82.1%;  Score 32;  DB 16;  Length 352;
  Best Local Similarity   55.6%;  Pred. No. 3e+02;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :||| :||:
Db        330 RLYEEIGMH 338


RESULT 15
Q8RBB7
ID   Q8RBB7      PRELIMINARY;      PRT;   365 AA.
AC   Q8RBB7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Bacterial cell division membrane protein.
GN   FTSW OR TTE0905.
OS   Thermoanaerobacter tengcongensis.
```

```
OC   Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC   Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4 / JCM 11007;
RX   MEDLINE=21992816; PubMed=11997336;
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
DR   EMBL; AE013057; AAM24161.1; -.
DR   GO; GO:0000910; P:cytokinesis; IEA.
DR   InterPro; IPR001182; Cell_cycle.
DR   Pfam; PF01098; FTSW_RODA_SPOVE; 1.
DR   PROSITE; PS00428; FTSW_RODA_SPOVE; 1.
KW   Cell division; Complete proteome.
SQ   SEQUENCE   365 AA;  40320 MW;  0CCAEC254B1E81E2 CRC64;

  Query Match            82.1%;  Score 32;  DB 16;  Length 365;
  Best Local Similarity  44.4%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XLYENVGMX 9
              :::||:||:
Db        314 HIFENIGMT 322


Search completed: July 15, 2004, 07:30:42
Job time : 35 secs
```